

المملكة العربية السعودية
وزارة التعليم العالي
جامعة أم القرى
كلية التربية
قسم علم النفس

تأثير حجم العينة على قوة الاختبار الإحصائي

إعداد الطالب
محمد إبراهيم احمد الشاردي

إشراف سعادة الأستاذ الدكتور
ربيع سعيد طه
أستاذ الإحصاء والبحوث

رسالة مقدمة لقسم علم النفس في كلية التربية بجامعة أم القرى متطلب تكميلي للحصول على درجة الماجستير في علم النفس –تخصص إحصاء وبحوث

الفصل الدراسي الثاني
١٤٣٢هـ/١٤٣٣هـ

مستخلص الدراسة

عنوان هذه الدراسة تأثير حجم العينة على قوة الاختبار الإحصائي

حيث شملت عينة الدراسة عينات عشوائية من المفردات الإحصائية التي تم توليدها تراوحت من (١٠) مفـــردات إلى (٣٣٠) مفردة، حسب طبيعة الاختبار الإحصائي المستخدم. وروعي في الأوامر المعطاة لبرنامج (PASS11)، أن تتحقق في هـــذه البيانـــات مجموعة الافتراضات اللازمة لاختباري (ت) و(ف)،سواء فيما يتعلق بالاختيار العشوائي، مستوى القياس الفئوي، التوزيع الإعتدالي، تجانس التباين.

ومن خلال النتائج تم التوصل الى أنه مع زيادة حجم العينة نلاحظ ازدياد قوة اختبار (ت) لعينه واحدة فعندما كـــان حجـــم العينة(n=30) كانت قوة الاختبار تساوي (٠،٨٧٧) بينما انخفضت قيمة (Beta) الى (٠،١٢) ، أما في نتائج اختبار (ت) لعينتين مترابطتين كان حجم العينة (n=30) مناسباً جداً للحصول على قوة اختبار عالية (٠،٨٥٢) وقيمة (Beta) منخفضة (٠،١٤٨) ، وفي اختبار (ت) لعينتين مستقلتين كانت قوة الاختبار عاليه من (٠،٧٧٣ — ٠،٨٥٦) وقيمة (Beta) منخفضة (٠،٢٢٨- ٠،١٤٤) عندما كان حجم العينة من (٤٠-٥٠) ، كما لوحظ أنه مع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين في حالة عدم تساوي حجوم العينات ، وفي الحالات التي تماثل بيانات البحث الحالي فأن النتائج تشير الى أهمية أن حجوم العينات كبيرة في كـــلآ العينتين.

كما توصلت نتائج اختبار (ف) في حالة تساوي حجوم العينات الى قوة اختبار عالية تراوحت مـــن (٠،٧٥٤) الى (٠،٨٤٩) وقيمة (Beta) منخفضة تراوحت من (٠،١٥١) الى (٠,٢٤٦) عندما كان حجم العينة ($40 \leq n \leq 50$) ، فيما لوحظ أن هناك تأثير كبير لحجم العينة على قوة اختبار (ف) في حالة عدم تساوي حجوم العينات عندما كانت ($150 \leq n \leq 180$) حيث كانـــت قـــوة الاختبار عالية تراوحت من (٠،٧٨٤) الى (٠،٨٥٦) وقيمة (Beta) منخفضة تراوحت من (٠،٣٢٩) الى (٠،٢١٦).

وبعد التوصل الى هذه النتائج تم اقتراح بعض التوصيات ومنها زيادة حجم العينة إلى العدد الذي يعطي قوة اختبار مناسبة والتي تقدر بحوالي (٠,٨) في مجال العلوم التربوية، ونقص في قيمة الخطأ من النوع الثاني والذي يقـــدر بحـــوالي (٠,٢) في مجـــال العلـــوم التربوية،وكذلك عدم المبالغة في زيادة حجم العينة عن الحد الذي يعطي قوة اختبار مناسبة،مع مراعاة أن يكون حجوم العينات مناسباً لنوع الاختبار الإحصائي المستخدم.

Abstract

The Effect Of Sample Size On The Statical Test Power

Where the study sample included a random sample of statistical individual that have been generated ranged between (10) items and (330) items, depending on the nature of the statistical test used. In orders given to the test (PASS11), it takes into accounts to realize in this data the set of assumptions needed to make test (T) and (F) with respect to random testing, categorical level of measurement , normal distribution and equinoctial variance.

Through the results it was concluded that with increasing sample size, noted that increasing of power of (t) test for one sample, when the sample size was (n=30) the test power was (0.877) while the value of (Beta) declined to (0.12), as the result of the test (T) for two interrelated samples the size of the sample (n=30) was very appropriate for high-power test (0852) and the value of (Beta) declined to (0.148), in the test (T) for two independent samples the power of the test was high (0.773 – 0.856) and the value of (Beta) is low (0.228-0.144) when the sample size between 40 – 50, also it is noted that with the increase of the size of the sample, the power of test (T) increases, for two independent samples in the case of unequal sample sizes, and in similar cases the results indicate to the importance of large sizes of sample must be in both samples.

Also the result of (F) test found that in case of equal sizes of the samples there will be high power of test between (0.754 to 0.849) and the value of (Beta) is low between (0.151 and 0.246) when the size of the sample is ($40 \leq n \leq 50$), as it is noted the great effect of the sample sizes on the power of the (F) test in case of unequal sample sizes when ($150 \leq n \leq 180$), where the high power test ranged between (0.784) and (0.856) and the value of (beta) was low between (0.329) and (0.216).

After reaching these results, some recommendations has been proposed including increasing the size of the sample to the number that gives the power of test to the suitable test estimated at about (0.8) in the field of educational sciences, and decrease the value of the error of the second type which is estimated at about (0.2) in the field of educational sciences, as well as the exaggerate in increasing the sample size of the limit which gives the power of an appropriate test taking into account that the size of the sample appropriate to the type of the used statistical test .

إهـــــداء

أُهدي نتاج عنائي، ثمرة جهدي وأبحاثي، شهد تعبي واشتغالي.

أُهديه أولاً وقبل كل شيء إلى من ربياني صغيرا .وتحملاني كبيرا..الى من أغدقا علي الغالي والنفيس سرورا.. من بهما بعد الله وصولي.. من اقبل مواطئ قدميهما فخراً وحبورا.. فأن كان جهد المقل يستحق الثناء فلهما بعد الله الثناء كله والفضل كله كثيرة ويسره، الى والدي يرحمه الله من ذكره مسكاً وكافورا / إبراهيم أحمد الشاردي ..

وإلى ذات الثلاث وصايا الى الغالية ولا غالي مثلها كبيراً كان أو صغيرا والدتي العزيزة أمدها الله بالصحة والعافية وقر عيني بشفائها

إلى زوجتي الغالية وشريكة تعبي واهتمامي ❖ أم إبراهيم ❖ ..

إلى أبنائي فلذات كبدي ،ثمرات قلبي العزيزة ،نجوم عالمي المنيرة

(وسن وإبراهيم وتالين)

على تحملهم وصبرهم علي طوال فترة الدراسة

إلى أخواي فخري وسندي ومعتمدي بعد الله في هذه الدنيا الأستاذ احمد إبراهيم الشاردي والأستاذ راشد إبراهيم الشاردي والى أبنائهم الأعزاء على نفسي .

إلى أخواتي وأبنائهم أشقاء روحي

وإلى فضيلة قاضي التمييز ورئيس محاكم القنفذة سابقاً الشيخ: محمد بن علي بن محمد الشاردي الذي أقبل رأسه احتراماً وتبجيلا ، من ملئ نفسي بحضوره مناقشتي زهواً وتشريفا .

إلى كل طالب علماً للفكر سفيرا

شكر وتقدير

الحمد لله العلي القدير والصلاة والسلام على نبينا محمد صلى الله عليه وسلم وعلـــى آلـــه وصحبه وبعد ..

أحمد الله الذي وفقني لإتمام هذه الرسالة ويسرني أن أتقدم بعميـــق شـــكري وتقـــديري لأستاذي الفاضل سعادة **الأستاذ الدكتور/ ربيع بن سعيد طه** على ما أفادني به من علمه الواسع ، وما قدمه من آراء وإرشادات قيمة وإهتمام خلال مراحل الدراسة والتي كان لها الأثر الكـــبير على إنجاز هذا العمل .

كما أتوجه بخالص الشكر والتقدير إلى كل من سعادة **الدكتور/ وجيه عبدالله مصطفى** ، وسعادة **الدكتور/ محمد معزوز قماري** على قبولهما مناقشة هذه الرسالة ، وما ســـيقدمانه مـــن توجيهات تساهم في إثراء هذه الدراسة، كما لا يفوتني أن أتقدم بالشكر لجميع الأساتذة الـــذين تكرموا بتحكيم خطة الدراسة وما قدموه لي من نصح وإرشاد ، كما لا يفوتني أن أشكر جميـــع من ساهم وتعاون في إنجاح هذه الدراسة سائلاً الله – عز وجل – التوفيق والسداد للجميع .

الباحث

قائمة المحتويات

قائمة الجداول

قائمة الأشكال

الفصل الأول

مدخل إلى مشكلة الدراسة

مقدمـــــــــــــــة .

مشكلة الدراسة وتساؤلاتها .

أهميـــــــة الدراســـــة.

أهـــداف الدراســـــة .

حـــدود الدراســـــة .

مصطلحـــات الدراســـة .

مقدمة Introduction :

الحمد لله رب العالمين والصلاة والسلام على اشرف الأنبياء والمرسلين سيدنا محمد وعلى آله وصحبة وسلم وبعد ..

إن ما نشاهده من الأعداد الضخمة للدراسات المنشورة عربياً وأجنبيا وازديادها المطرد ونظراً لهذا الازدياد والتوسع فإن التوجه لمراجعة الأبحاث أصبح محبباً في كثير من المجالات التربوية ومن أهم التوجهات هي التي تركزت على طرق البحث والمعالجات الإحصائية المستخدمة وذلك لان مجال الإحصاء مجال دائم التغير والنمو والتطور ويخضع إلى التحسينات والإضافات بصورة دائمة ومستمرة،ومن المعلوم أن جميع الأدوات الإحصائية ذات الطابع الاستدلالي تتضمن مجموعة من الافتراضات الأساسية تكون الأداة الإحصائية صادقة في حالة توفر الافتراضات ،و تكون الأداة الإحصائية غير صادقة في حالة عدم توفر الافتراضات .

وقد تطور علم الإحصاء كنتيجة طبيعية لتطور المعارف الأخرى بسبب زيادة الاهتمام بالبحث وعملية اتخاذ القرارات،وعليه ظهرت العديد من النظريات والأساليب الحديثة (تشاو،١٩٩٠م) التي من أبرزها أساليب فحص الفرضيات التي تشكل الأساس لاتخاذ القرارات باستخدام نوعين من الدلالات أحدهما إحصائية، والثانية عملية.حيث أن الدلالة الإحصائية تعني احتمالية رفض الفرض الصفري Ho باستخدام الأساليب الإحصائية،بينما يقصد بالدلالة العملية أن تكون الفروق الإحصائية أو العلاقات بين المتغيرات كبيرة إلى درجة تبرز عملية الأخذ بنتائجها.

حيث يُعرف علم الإحصاء بأنه ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وتلخيصها وعرضها وتحليلها وذلك للوصول إلى نتائج مقبولة تؤدي إلى اتخاذ قرارات سليمة (بري وهندي والحسيني،١٩٩٨م).

ويشير كلاً من (Hays,1973)، (Cohen,1977) إلى إن الدلالة الإحصائية شرط ضروري لصناعة قرار تربوي أو نفسي ولكن ليس كافياً،فالكفاية تتحقق بحساب قوة العلاقة بين المتغير المستقل والمتغير التابع، وقوة العلاقة هي ما يسمى بالدلالة العملية الصياد(١٩٨٨م).

وأكد سيجل (١٩٥٦م)على أن أصعب مرحلة تواجه الباحث في بحثه هي مرحلة التحليل الإحصائي ،وذلك لكثرة تعداد أنواع الطرق الإحصائية. ولأهميتها عمد سيجل إلى تحديد أربعة معايير لابد أن يأخذها الباحث بعين الاعتبار عند اختياره للطريقة الإحصائية وأشار إليها (توفيق،١٩٨٥م) هي:

١- طبيعة توزيع متغيرات الدراسة في المجتمع الذي اختيرت منه العينة.

٢- نوعية مستوى القياس المستخدم.

٣- تصميم البحث من حيث عدد العينات ونوعها وحجمها.

٤- قوة الاختبار.

ويلاحظ أن بعض الباحثين يعتمدون في تقدير نتائجهم على الدلالة الإحصائية دون محاولة الكشف عن مقدار العلاقة القائمة بين المتغيرين وتصبح هناك مغالاة في تفسير النتائج اعتماداً على مستوى الدلالة على الرغم من أنه ربما لا تكون لها قيمة من الناحية التطبيقية أو العملية،ولذلك فإذا وجد الباحث أن القيمة دالة إحصائيا،فمعنى ذلك أن المتغير المستقل له تأثيره على المتغير التابع،ولكنه لا يدل على حجم التأثير أو درجة العلاقة بين المتغيرين،وربما كانت الدلالة الإحصائية هنا لا تعني وجود علاقة قوية بين المتغيرين(الشربيني،٢٠٠٧م).

وبناءً عليه فإن إهمال الدلالة العملية ((حجم التأثير)) وأهميتها يقلل من أهمية نتائج الدراسة وقد يؤدي إلى التضليل،فالأساس هو تقدير الدلالة العملية في البحث التربوي،خصوصاً إذا علمنا أن قوة الاختبار الإحصائية تكمن في علاقتها بالدلالة الإحصائية والدلالة العملية.حيث أشار(Wambold,1983) إلى أن هذه العلاقة تأخذ إحدى الصور التالية:

١. إذا كانت هناك دلالة إحصائية ودلالة عملية وكانت قوة الاختبار عالية فيكون القرار سليماً.

٢. إذا كانت هناك دلالة إحصائية دون دلالة عملية وقوة الاختبار منخفضة فإن نتائج القرار تكون مضللة.

٣. إذا كانت هناك دلالة عملية دون دلالة إحصائية وقوة الاختبار منخفضة (في هذه الحالة لا يؤخذ بالقرار المتخذ).

٤. إذا لم يكن هناك دلالة إحصائية وكانت قوة الاختبار عالية عندها تكون الدلالة العملية عالية،(وفي هذه الحالة يكون القرار سليماً) (الصائغ، ١٤١٧هـ).

وإشارة إلى ما سبق من الدراسات التي تناولت الاختبارات الإحصائية مثل دراسة البابطين(٢٠٠٢)، والصائغ(١٤١٧)، ونور(١٤١٣)، والنجار(١٤١١) فجميعها أوصت بالاهتمام بمعرفة الدلالة العملية (حجم الأثر) وحجم العينة حيث أنها لم تبحث من قبل في جامعة أم القرى ، أما على الجانب العربي فقد تناولها الصياد(١٩٨٨،١٩٨٩)وأوصى بضرورة البحث في واقع قوة الاختبار الإحصائي المستخدم في البحوث التربوية والنفسية كبعد هام وأساسي لمستوى الدقة في البحث،وعامل هام من عوامل التأثر على الدلالة العملية ومعطياتها في مجالات الإحصاء التطبيقي.وتؤكد دراسة (Huston,1993) في نتائجها على أن الباحثين الذين لا يعتمدون على الدلالة الإحصائية وحدها ويلجأون الى حساب قوة الاختبارات الإحصائية وحجم التأثير هم الذين يساهمون في تحسين نوعية البحوث.

وبناءً عليه فإن تحليل القوة الإحصائية وتقدير حجم التأثير في الدراسات العلمية التي سبقت هذه الدراسة يوضح واقع البحوث التربوية وجوانب قصورها،ومن هذا المنطلق يرى الباحث أن هذه المحاولة لها أهميتها على حد علمه في بعض الأساليب الإحصائية شائعة الاستخدام في الدراسات التربوية،مما يسهل للباحثين التعرف على أهمية حجم التأثير لمتغيراتهم المتعددة والتي تتأثر بعوامل حجم العينة وقوة الاختبار الإحصائي ، وهو ماتم التركيز عليه في هذه الدراسة .

مشكلة الدراسة وتساؤلاتها The problem of Study :

في الأبحاث التربوية عادةً يتم اللجوء إلى تقنيات الإحصاء الاستدلالي وذلك من اجل تحديد حجم الأثر أو الفرق أو العلاقة بين المتغيرات والتأكد من دلالتها الإحصائية، و التي ترضخ لعلامات كثيرة كحجم العينة وعدد المتغيرات وقوة الاختبار ومستوى الدلالة، ونلاحظ إن كثير من الدراسات المنشورة أهملت العديد من الشروط العلمية التي يجب مراعاتها عند القيام بأي بحث علمي . وقد أكد الصياد على أهمية قوة الاختبار كعامل مؤثر في كل من حجم العينة وحجم التأثير ((الدلالة العملية)) (الصياد،١٩٨٩) ومما سبق يمكن تحديد مشكلة الدراسة في التساؤلات التالية :

١- ما تأثير حجم العينة على قوة اختبار (ت) لعينة واحدة؟

٢- ما تأثير حجم العينة على قوة اختبار (ت) لعينتين مترابطتين؟

٣- ما تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين؟

٤- ما تأثير حجم العينة على قوة اختبار تحليل التباين الأحادي (ف) للمقارنة بين المتوسطات الحسابية؟

أهمية الدراسة Importance of the study :

تقوم الدراسة الحالية بتوضيح تأثير حجم العينة وكذلك تأثير عدد المتغيرات على قوة اختبار (ت) و (ف) ويمكن تحديد أهمية البحث في التالي :

١- يمكن للبحث الحالي أن يقدم للباحثين المعلومات اللازمة للتعامل مع ظاهرة أغفلها الباحثون وهي الاهتمام بحجم العينة وعدد المتغيرات .

٢- التعرف على أهمية قوة الاختبار الإحصائي وعلاقته بحجم العينة والمتغيرات.

٣- الاهتمام بأهم العوامل المؤثرة على قوة الاختبار الإحصائي والتي من المتوقع وجودها في الرسائل العلمية المنشورة .

٤- التعرف على حجم العينات المناسب لنوع الاختبار الإحصائي المستخدم في الدراسات التربوية.

أهداف الدراسة Purpose The Study :

الهدف من الدراسة الحالية هي محاولة لإلقاء الضوء على أهم الطرق الإحصائية لحساب وتقدير حجم العينة لبعض ألاختبارات الإحصائية وتوضيح طريقة التطبيق وتقويم النتائج وذلك من اجل الاعتماد عليها في اتخاذ قرارات دقيقه تجاه الظاهرة المدروسة وبناءً على الهدف العام فقد تفرعت منه الأهداف التالية :-

١. التعرف على أهم الطرق الإحصائية لحساب تأثير حجم العينة على قوة الاختبار الإحصائي في بعض الأساليب الإحصائية شائعة الاستخدام في الدراسات التربوية.
٢. التعرف على واقع تأثير حجم العينة على قوة اختبار(ت) لعينة واحد، ولعينتين مترابطتين، وكذلك لعينتين مستقلتين.
٣. التعرف على واقع تأثير حجم العينة على قوة اختبار تحليل التباين الأحادي(ف) للمقارنة بين المتوسطات الحسابية.

حدود الدراسة Limitations of the Study :

تتحدد الدراسة الحالية بدراسة حجم العينة على قوة الاختبار الإحصائي في أسلوبي (ت) ، (ف) .

تعريف بمصطلحات الدراسة Definitions of the Study :

الأساليب الإحصائية :

هي تلك الطرق الإحصائية التي تهتم بالبيانات المجموعة عن ظاهرة ما ،والتي يستخدمها الباحث بغرض تبويب هذه البيانات وتحليلها مما يساعد على استخلاص النتائج منها واتخاذ القرارات. (النجار ١٤١١هـ)

المتغير :variable

يعرف المتغير بأنه الخاصية التي تأخذ قيماً مختلفة للأفراد المختلفين في المجموعة قيد الدراسة(المنيزل،٢٠٠٠).

حجم العينة :

عامل هام ومؤثر بصورة طردية في قوة الاختبار فزيادة قوة الاختبار تنتج من ازدياد حجم العينة.

ويقصد بحجم العينة عدد أفراد عينة البحث الممثلة للمجتمع الأصلي .

قوة الاختبار:Power of the test

هي قدرة الاختبار على رفض الفرض الصفري ،عندما يكون في حقيقة الأمر خاطئا ، وتكون تلك القوة في صورة احتمال تعتمد على احتمال ارتكاب الخطأ نمط اثنين (الشربيني ، ٢٠٠٧م) .

حجم التأثير :Effect Size

هو عبارة عن قيمة تشير إلى درجة العلاقة بين متغيرات الدراسة ضمن مجتمع محدد مسبقاً هو مجتمع الدراسة بغض النظر عن الأسلوب الإحصائي المستخدم للتعبير عن ذلك الأثر(نصار،٢٠٠٢).

مستوى الدلالة :Leve of Significance

وهو احتمال رفض الفرض الصفري وهو صحيح،ويرمز له بالرمزα.

الخطأ من النوع الثاني:Type II Error

يعرف باسم بيتا(β) وهو احتمال قبول الفرضية الصفرية وهي خاطئة .

اختبار(ت) :T- test

اختبار(ت) أو (ت الطالب) هو اختبار إحصائي لدلالة الفرق بين متوسطي عينتين مستقلتين أو مترابطتين . (فرج،١٩٩٦).

اختبار " ت" عندما يكون متحررا :Liberally في حالة عدم تساوي أحجام العينات بحيث يكون التصميم الإحصائي غير متزن وعندما يكون التباين الأصغر مع المجموعة الأكبر حجما.

اختبار " ت" عندما يكون متحفظا :Conservative في حالة عدم تساوي أحجام العينات وبحيث يكون التصميم الإحصائي غير متزن وعندما يكون التباين الأكبر مع المجموعة الأكبر حجما.

تحليل التباين الأحادي الاتجاه: One-Way Anova

هو أسلوب إحصائي يستخدم لاختبار الفروق بين عدد من المجموعات المستقلة تزيد عن مجموعتين في متغير تابع واحد .

الفصل الثاني

(الإطار النظري والدراسات السابقة)

أولاً : الإطار النظري

- **المبحث الأول:**

أولا:قوة الاختبار Power of the Test.

ثانيا:حجم التأثير(الدلالة العملية) Effect Size

- **المبحث الثاني:**

أولا: اختبار " ت " (t-test).

ثانيا: واختبار تحليل التباين(ANOVA)

ثانياً : بحوث ودراسات سابقة

الإطار النظري

المبحث الأول

أولاً:قوة الاختبار Power of the Test:

قوة الاختبار تعني رفض الفرض الصفري عندما يكون خاطئاً بالفعل ، وهو يـــدل علـــى المساحة الواقعة تحت المنحنى وبتحديد في منطقة الرفض.ويـــذكر الـــشربيني(٢٠٠٧م) إن قـــوة الاختبار في البحوث الإنسانية بخاصة تكون مقبولة إذا كانت تتراوح مابين (٠,٤٠-٠,٦٠).كما أشار علام (٢٠١٠م) إلى أن قوة الاختبار تتراوح مابين (صفر-١) نظراً لأنها احتمالية . في حين اعتبر كوهين (٠,٥٠) كحد أدنى لقوة الاختبار.أما الصياد فيرى أن تكون قوة الاختبار(٠,٨٠) ،وتفضل الصائغ أن تكون قوة الاختبار (٠,٩٠).

قوة الاختبار والخطأ من النوع الثانيβ:

إن الرفض يكون على صورة احتمال تعتمد قيمته بشكل مباشر على احتمال ارتكاب الخطأ من النوع الثاني نظراً للعلاقة العكسية الموجودة بين قوة الاختبار(ρ) والخطأ من النوع الثاني (β) ولذلك يجب أن تكون β عند أقل مستوى ممكن .

العوامل المؤثرة على قوة الاختبار:

هناك مجموعة من العوامل المؤثرة على قوة الاختبار لعل أهمها مايلي:

١- حجم العينة .

٢- مستوى الدلالة α .

٣- علاقة القيمة الحقيقية للمجتمع بقيمته في الفرضية الصفرية .

٤- الاختبار بذيل والاختبار بذيلين .

٥- الاختبارات المعلمية والاختبارات اللامعلمية.

٦- حجم التأثير .

وقد اتفق على هذه العوامل المؤثرة في قوة الاختبار كلاً من (كوهن، ١٩٧٧م)،و(الشربيني، ٢٠٠٧م)، و(عودة والخليلي،٢٠٠٠م)، و(علام ،٢٠١٠م)و(الصائغ،١٤١٧هـ).

وفيما يلي شرح مبسط لكل عامل من هذه العوامل:

١- حجم العينة:

يرى كوهين (cohen)وهيس (Hays) إن العلاقة بين حجم العينة وقوة الاختبار علاقة مباشرة فالزيادة في حجم العينة تزيد من القوة الإحصائية مع ثبات العوامل الأخرى عندما يكون الفرض الصفرى غير صحيح ، ويكون القرار رفض الفرض الصفري .ويتأثر احتمال الخطأ من النوع الثاني (β) وقوة الاختبار باختلاف أو تباين المجتمع وحجم العينة ، حيث أن كلما قل تباين المجتمع أو زاد حجم العينة قل احتمال الخطأ من النوع الثاني (β) وزادت قوة الاختبار ، ويرجع التباين في حجم المجتمع إلى انخفاض ثبات أدوات القياس أو عدم اتساق الدرجات المستمدة منها (الصائغ،١٤١٧).

كما بين الصياد (١٩٨٩م) جداول لتحديد حجم العينة بمعرفة مدخل كوهين حيث صنف الجداول إلى صورتين :

١- عندما يكون الاختبار الإحصائي معلوماً من قبل الباحث للاختبارات التالية شائعة الاستخدام (t ,f ,x^2 ,r) .

٢- تحديد حجم العينة في ضوء عدم معرفة الاختبار الإحصائي المستخدم .

وبما أن قوة الاختبار تزداد بزيادة حجم العينة ،فزيادة حجم العينة قد يفيد في حالات معينة وهي :

١- وجود متغيرات غير مضبوطة بأي طريقة من طرق ضبط المتغيرات .

٢- توقع إعادة تقسيم المجموعة الكلية إلى مجموعات جزئية حسب مستويات المتغيرات المستقلة.

٣- عندما لا يكون المجتمع متجانساً.

٤- عندما يكون ثبات المقياس للمتغير التابع منخفضاً، مما يعني وجود أخطاء في المقياس .

وعند تحديد حجم العينة يأخذ الباحث بعين الاعتبار عدة أمور أهمها :

أ- طبيعة المجتمع الأصلي (أي هل هو متجانس أو متباين) .

ب- طريقة اختيار العينة .

ج- درجة الدقة المطلوبة .

كما أضاف كوهين (cohen,1977) ثلاثة عوامل :

د- حجم الأثر الأولي المحدد من قبل الباحث منسوباً للانحراف والمحايدة.

هـ- قوة الاختبار الإحصائي المستخدم في تحليل النتائج .

و- اتجاه الاختبار ،بمعنى هل هو(بذيل أو بذيلين) .

وبناء على العوامل السابقة الذكر وضع كوهين (cohen) جداول لحساب حجم العينة وتحديدها و التي سميت بمدخل كوهين لتحديد حجم العينة أو مدخل اختبارات الفروض (التجريبي وشبه التجريبي) في تحديد حجم العينة .

٢- مستوى الدلالة α:

يتناسب مستوى الدلالة تناسباً طردياً مع قوة الاختبار ، فإذا قل مستوى الدلالة من (.05) على سبيل المثال إلى (.01) فإن حدود الرفض (المنطقة الحرجة)تتغير من :(Z=±1.969 إلى Z=±2.58) وهذا يعني ازدياد صعوبة رفض الفرضية الصفرية، بمعنى أن احتمال ارتكاب الخطأ من النوع الأول يقل،لكن في الوقت نفسه يزداد احتمال ارتكاب الخطأ من النوع الثاني الأخر الذي يخفض من قوة الاختبار (ابوزينة ورفاقة ،٢٠٠٧).

أن المعيار الأساسي لاختبار الفروض الإحصائية من حيث اختبار مستوى الدلالة هو قوة الاختبار(ρ) بحيث يجب أن تكون قيمته على الأقل أكبر من حد الصدفة ٠,٥٠ ،في حين كوهن يرى أن تكون ٠,٨٠ كحد أعلى لتعامل الباحث مع المتغيرات غير الثابتة دائماً ، ويكون الخطأ من النوع الثاني β =٠,٢٠ ويمكن تحديد قوة الاختبار قبل إجراء الدراسة لمعرفة أي الاختبارات الإحصائية أكثر قوة للتعامل معه (الصياد،١٩٨٩).

٣- علاقة القيمة الحقيقية للمجتمع بقيمته في الفرضية الصفرية:

اتفق كلا من علام (٢٠١٠م)،وعودة والخليلي (٢٠٠٠م) على انه كلما زاد الاختلاف بين المتوسط الحقيقي (الأصلي للمجتمع) والمتوسط الفرضي زادت قوة الاختبار وقل الخطأ(β).

٤- الاختبار بذيل والاختبار بذيلين:

قوة الاختبار تزداد في حالة تحقق الباحث من صحة الفرض الصفري باختبار إحصائي ذي طرف واحد . فالباحث عندما يحدد المتوسط في ضوء الفرض الصفري ،وكان الافتراض صحيحاً، فأن الخطأ من النوع الثاني أقل مما لو استخدم اختبار بطرفين.

٥- الاختبارات المعلمية والاختبارات اللامعلمية :

إن الاختبارات المعلمية تعتبر أكثر قوة من الاختبارات اللامعلمية وذلك لأي قيمة معينة من قيم حجم العينة فاحتمالية الوقوع في الخطأ من النوع الثانيβ يكون أقل في حالة الاختبارات الإحصائية للفرض الصفري والتي تفترض اعتدالية التوزيع للمجتمعات وتساوي التباين حيث إن الأساليب المعلميه تراعي القيم الأصلية للدرجات وهذا يجعلها أكثر دقة في التوصل للاحتمالية عندما تتحقق الفروض. والأساليب الإحصائية المستخدمة لاختبار الفروض الصفرية Ho تختلف في مدى قوتها، فالتي ترفض الفرض الصفري تسمى أساليب أكثر قوة عن غيرها،أما الأساليب الأقل قوة فيفضل استخدامها في البحوث ذات العينات الكبيرة. ولذلك تعتبر قوة الاختبارات الإحصائية من أهم محكات اختيار النموذج الإحصائي المناسب .

٦- حجم التأثير :

أن قوة الاختبار تزداد كلما ازداد حجم الأثر ، وتقل قوة الاختبار كلما قل حجم الأثر، ابوزينة ورفاقة (٢٠٠٧م).وتضيف الصائغ (١٤١٧ه) انه كلما زاد حجم التأثير قل حجم العينة اللازم للوصول لقيمة معينة دالة إحصائياً ، ويمكن إيجاد قوة الاختبار بالاعتماد على حجم التأثير. وقد وضع كوهين عدة معايير للدلالة العملية في الاختبارات الإحصائية (f،t) على النحو التالي:

جدول رقم (١)

معايير الدلالة العملية في الاختبارات الإحصائية (t,f)

في حالة اختبار (t)
عندما تكون : d = ٠,٢٠ يكون حجم التأثير صغيراً
عندما تكون : d = ٠,٥٠ يكون حجم التأثير متوسطاً
عندما تكون : d = ٠,٨٠ يكون حجم التأثير كبيراً
في حالة اختبار (f)
عندما تكون : F = ٠,١٠ يكون حجم التأثير صغيراً
عندما تكون : F = ٠,٢٥ يكون حجم التأثير متوسطاً
عندما تكون : F = ٠,٤٠ يكون حجم التأثير كبيراً

وسوف يتم تناول هذا الجانب بالتفصيل في الجزء الثاني من المبحث الأول.

ثانياً: حجم التأثير (الدلالة العملية):

في بداية العشرينات تعرض لها بيرسون ،وتم تفسير مربع ايتا (η^2) عن طريق فيشر١٩٢٥م مع تحليل التباين ، وحددها كيللي ١٩٣٥م لتقييم المجتمع (ω^2) ،وفي بداية الأربعينات عُممت في كتب الإحصاء والقياس بمسمى حديث (حجم التأثير) أو (قوة الارتباط)، وتم تطويرها حديثاً بواسطة هيس ١٩٦٣م للمقياس (ω^2)، وفي بداية الستينات ظهر تطبيقها في البحوث التربوية عن طريق شوتز ١٩٦٦م ولكنه أهمل تطبيقها لفترة (١١-١٣عاماً) حتى عام ١٩٧٧-١٩٧٩م تقريباً ، بدأ الاهتمام بتطبيقها في الدراسات المعاصرة وتحديداً في مجال التربية وعلم النفس ومنذ عام ١٩٩٤ والجمعية النفسية الأمريكية(APA)تؤكد على أهمية الإشارة إلى قيمة حجم الأثر في التقارير البحثية وذلك في ضوء تأثر نتائج فحص الدلالة الإحصائية للبيانات بحجم العينة، وفي عام (٢٠٠١) تم التأكيد على نفس النقطة من خلال النسخة الخامسة من دليل نشر(Publication Manual) الجمعية النفسية الأمريكية (APA) لما لذلك من أهمية في توضيح

النتائج البحثية من هنا فإن العديد من المجلات البارزة في مجال علم النفس والعلوم التربوية أو العلوم الإنسانية تشترط على الباحثين تضمين قيمة حجم الأثر في أبحاثهم(الصائغ،١٤١٧).

فحجم الأثر مرتبطة بدلالة الفرق الإحصائي وتكمن في أهميتها العملية والتطبيقية والتفسيرية للقرار المتخذ ، فمهمتها مناقشة نتائج البحث ثم اتخاذ قرار عملي يطبق واقعياً دون تكلفة أو جهد للموازنة بين المكاسب والخسارة بصورة عملية .

لذلك فحجم الأثر يعتبر الوجه الآخر للدلالة لأنها تتعلق بالمزايا الفعلية الناتجة مـــن تقيـــيم الفرق الإحصائي المحسوب في اختبار دلالة الفرق .و قيمة حجم الأثر تشير إلى أي درجة يمكـــن التنبؤ بالمتغير التابع أو تفسيره من خلال المتغير المستقل (Mahadevan,2000).

واتفق كلا من كوهن (Cohen)، وهيس (Hays) ،وماكنمارا (Mcnamara) وآخرون على ضرورة حساب قوة العلاقة بين المتغير التابع والمتغير المستقل وأُطلق عليها اســـم (الدلالـــة العملية). ويطلق عليها حديثاً (حجم التأثير) ،أو قوة الارتباط الإحصائي.نظراً لان تقدير قـــوة الارتباط لها أهميتها في تقييم نتائج الدراسات فهي وسيلة للتطوير والتوجيـــه، بالإضـــافة إلى أن حسابها يقلل من الوقوع في أخطاء القرارات البحثية.

ويذكر الصياد (١٩٨٨م) عن حجم التأثير انه مؤشر لمدى قدرتنا على استخدام النتـــائج تفسيراً أو تطبيقاً، إنه كم التباين الذي أمكن تفسيره للمتغير التابع حينما اعتبرنا المتغير المستقل في علاقة معه أو مؤثر عليه.وقد أشار هس (Haase,1982) الى إن مشكلة تقييم الدلالة العمليـــة والدلالة الإحصائية لها تاريخ طويل بعض الشيء.

ويضيف الصياد (١٩٨٨)أن الدلالة الإحصائية لا تكفي لاتخاذ قرار نفسي أو تربوي لأنها شرط ضروري من شروط اتخاذ القرار ولكنها ليست كافية وحدها، ذلك أن الكفاية تتحقق فقط عندما نحسب قوة العلاقة بين المتغير المستقل والمتغير التابع ،وهي تُحدد حسب طبيعة الأشياء محل البحث ،وتحكمها القيم السائدة في المجتمع .

ويذكر نصار(٢٠٠٢) إن حجم الأثر هو عبارة عن أسلوب إحصائي يشير إلى الدلالة العملية للنتائج الإحصائية . ويؤكد علام (٢٠١٠م) على الباحث أن يضع في اعتباره أهمية الدلالة العملية أو التطبيقية والدلالة التفسيرية ،لأنها مترابطة ويكمل بعضها الآخر.

فحجم الأثر يعتبر مكمل للدلالة الإحصائية ومقياس إضافي يجب أن يُسلم به الباحث التربوي، ويستعين به بعد الاستنتاج الإحصائي لاختبار دلالة الفرق لاتخاذ قرار صحيح .

استخدام حجم الاثر في الدراسات النفسيه :

يشير كل من شتينبرج وثيسن (steinberg & thissen,2006)،والعبد القادر (١٤٢٩هـ) إلى أن هناك ثلاثة استخدامات مختلفة لأحجام التأثير في الدراسات النفسية:

١- استخدام حجم التأثير في تحليل القوة :

" وصف كوهن مجموعة من مؤشرات حجم التأثير التي لها دور الوسيط في حساب قوة رفض أو استبعاد الفرض الصفري في العديد من التصميمات التجريبية . وقد أكد كوهن وبكل وضوح إلى أن السبب وراء صياغة تلك الإحصاءات المعيارية هو " عدم أمكانية تجهيز وبناء مجموعة من جداول القوة الإحصائية Power Tables بالنسبة لكل وحدة قياس جديدة "وتمثل المقاييس المعيارية لحجم تأثير المجتمع الإحصائي قاعدة أساسية للجداول والعمليات الحسابية التي استخدمها كوهن في حساب القوة فضلاً عن ذلك قدم كوهن بعض القيم لتصنيف مؤشرات حجم التأثير إلى "صغير ومتوسط وكبير " مع الإشارة إلى بعض المبررات بشأن عدم استخدامها بشكل مباشر دون وجود خبرة في التعامل مع حجم التأثير في الدراسات النفسية . وقد هدف كوهن من ذلك التصنيف إلى تزويد الباحثين ومحللي البيانات بقيم عددية (اختيارية) لأحجام التأثير اللازمة لإجراء تحليل القوة .

٢- استخدام حجم التأثير في صياغة الأبحاث :

يلعب التحليل البعدي دوراً هاماً في دمج مجموعة من أحجام التأثير في صورة تقدير وحيد Single Estimate لحجم التأثير الشامل ، وذلك من أجل توضيح طبيعة توزيع أحجام التأثير أو بناء النموذج الاستطلاعي للتنوع في حجم التأثير . وفي بعض الحالات يكون حجم التأثير

المستخدم في التحليل البعدي هو ذاته المستخدم في تحليل القوة . وفي أطار التحليل البعدي Meta Analysis لابد من وجود شرط معين الاستخدام حجم التأثير ، وهو ارتباط تقديرات هذا الأخير بتقدير الدقة Precision Estimate .

٣- استخدام حجم التأثير في تقرير الأبحاث :

" هناك مبدأ عام يجب إتباعه أنه لابد من تزويد القارئ ليس فقط بالمعلومات عن الدلالة الإحصائية ولكن أيضاً بمعلومات أخرى كافية من أجل قياس حجم التأثير الملحوظ للمعالجة الإحصائية " (APA 2001).

ولقد أورد كل من باركر وهاجان (Parker and Hagan 2007) بعض المميزات لحجم التأثير :

١- مقياس موضوعي لقوة التدخل Intervention Strength (وذلك في حالة صدق وقوة التصميم التجريبي).

٢- مؤشر له تدريج مستمر Continous Scale ، وبالتالي يدعم القرارات التدريجية في المعالجة الإحصائية .

٣- مقياس دقيق خصوصاً في حالة النتائج الصغيرة وغير الواضحة .

٤- ملخص موضوعي في حالة عدم الاتفاق على الأحكام البصرية بشأن الرسوم البيانية .

٥- طريقة للمقارنة بين النجاحات المختلفة للتدخل عبر أبحاث دراسة الحالة ، سواء كان التحليل آنياً أو بعدياً .

٦- مقياس معتمد في أبحاث دراسة الحالة بالمقارنة مع التقاليد البحثية الأخرى .

٧- طريقة فعالة لتوثيق النتائج .

٨- مقياس جيد لاستقلالية النتائج ، خصوصاً عند استخدام فترات الثقة .

حجم التأثير وعلاقته بحجم العينة:

يذكران سنايدر ولوزن (Snyder and Lawson 1992) ان حجم العينه يعتبر العامل الأساس في التقرير فيما إذا كانت نتائج أي دراسه داله أو غير داله إحصائياً. وكما يشير كارج و يسون و ميتز (Craig, Eison & Metz 1976) إنه كلما كبر حجم عينة الدراسه كلما زاد إحتمال الحصول على نتائج داله إحصائياً حتى ولو كان هناك إرتباط ضعيف ضمن مجتمع الدراسه بين المتغيرات المستقله والمتغيرات التابعه.

ويرى هيس(Haya1963) بأن الدلاله الإحصائيه لا تشير بالضروره إلى قوة العلاقه بين متغيرات الدراسه بل هي في أحسن الأحوال تشير إلى بعض الإرتباط بين هذه المتغيرات دون تحديد مدى أهمية ذلك الإرتباط. هذا و أضاف هيس في هذا الخصوص أن العامل الأهم في الحصول على نتائج داله إحصائياً هو حجم العينه وأنه من المهم التحديد الدقيق لمفهوم قوة العلاقه الإحصائيه، ويمكن تفسير رأي هيس على أساس إعتبار حجم العينه مضللاً في كثير من الأحيان بحيث أن زيادة حجم العينه يؤدي إلى الحصول على نتائج داله إحصائياً مما يضخم من حجم العلاقه بين المتغيرات المستقله والمتغيرات التابعه ومما يحد من القدره على تحديد حجم العلاقه الحقيقيه بين هذه المتغيرات.

وقدم ماكسويل و ديلني (Maxwell and Delaney 2004) معادله (قيمة الإحصائي = حجم الأثر × حجم الدراسه)، حيث تهدف لتوضيح العلاقه بين قيمة الإحصائي (z ، t ، f، ... ألخ) وحجم عينه الدراسه وحجم الأثر.

المؤشرات الإحصائية المستخدمة للدلالة على حجم الأثر في الأساليب الإحصائية المختلفة:

يرى كلاً من أبو هاشم (٢٠٠٥)،وماكسويل وديلني (٢٠٠٤)،ونصار(٢٠٠٢) ، انه يمكن تقسيم المؤشرات التي تدل على مقدار حجم الأثر في الدراسات المختلفة إلى نوعين هما:

(١) المؤشرات التي تدل على مقدار الأثر الذي يُحدثه المتغير المستقل في المتغير التابع (Measures of effect size) وتُعرف هذه المؤشرات أيضاً بأنها الفرق بين المتوسطات (متوسط المجموعة

التجريبية ومتوسط المجموعة الضابطة في الدراسات التجريبية) بوحدات معياريه (عوده والخليلي ، ٢٠٠٠).

وتستخدم هذه المؤشرات في الدراسات التي تعتمد على الأساليب الإحصائية التي تفحص الفروق بين المتوسطات، كما هو الحال في اختبار "ت" سواء للعينات المستقلة أو العينات المترابطة وأشهر هذه المؤشرات هو المؤشر الذي قدمه كوهين (Cohen,1988) ويطلق عليه الرمز " d".

وهناك مؤشرات أُخرى مثل مؤشر كوهن f (Cohen,1988)، ومؤشر Δ(Class,1976)، ومؤشر هدجز g(Hedges,1981).

(٢) المؤشرات التي تدل على قوة العلاقة أو الارتباط بين متغيرات الدراسة (Measures of association) وتستخدم هذه المؤشرات في الأساليب الإحصائية التي تعتمد على فحص التباين حيث تشير إلى نسبة التباين الذي يمكن تفسيره في المتغير التابع عن طريق تباين المتغير أو المتغيرات المستقلة. ومن أمثلتها مؤشر مربع إيتا لبيرسون η^2 (Pearson,1905)،ومؤشر مربع أوميغا لهاس ω^2 (Hays,1963)،ومؤشر مربع أبيلسون لكيلي ε^2 (Kelley,1935).

وهناك العديد من المؤشرات والتي تستخدم في الاساليب الإحصائية المختلفة وسوف يستعرض الباحث بعض انواع المؤشرات التي تُستخدم في الاساليب الإحصائية وبالتحديد في اختبار " ت " t-test للعينات المستقلة والمترابطة واختبار تحليل التباين الأحادي " ف " ANOVA:

١- المؤشرات التي تُستخدم للدلاله على حجم الأثر في اختبار " ت " للعينات المترابطة والمستقلة:

يختلف حساب حجم التأثير بأختلاف الإحصائي المستخدم في فحص الفرضية الصفرية،وقد أوجد كوهن Cohenطُرقاً مختلفة لحساب حجم التأثير لـ الإحصائي ت (t) والإحصائي ف (f) (ابراهيم،١٩٩٨).

- مؤشر كوهين (Cohen's d):

$$\text{حجم التأثير} = \frac{\text{متوسط المجموعة التجريبية} - \text{متوسط المجموعة الضابطة}}{\text{الانحراف المعياري}}$$

حيث تقوم فكرة هذا المؤشر الذي قدمه كوهين (١٩٨٨) على إستخراج الفرق بين متوسط المجموعة التجريبية والمجموعة الضابطة ومن ثم قسم هذا الفرق على الإنحراف المعياري لإحدى المجموعتين على اعتبار أن كلاّ المجموعتين متجانستين من حيث التباين وهذا هو احد الإفتراضات (Assumptions)التي يفضل توفرها عند استخدام اختبار " ت ".

ومن الأمثلة المعروفة للباحثين في مجال التربية وعلم النفس حالة العينتين المستقلتين لاختبار (ت) مثل التصميم التجريبي الذي يتضمن مجموعة ضابطة Group Control وأخرى تجريبية Experirnentol Group،ويمكن حساب حجم التأثير من المعادلة :

(الضوي،٢٠٠٦)

$$ES = T\sqrt{\frac{1}{n_E} + \frac{1}{n_c}}$$

حيث أن:

ES=حجم التأثير

n_E=حجم المجموعة التجريبية

n_c= حجم المجموعة الضابطة

t = قيمة(ت) المحسوبه

ويضيف نصار(٢٠٠٢م) انه عندما تكون العينتين التجريبية والضابطة غير متساوية الحجم (عدد أفراد إحدى المجموعتين أكبر من الأخرى) فإن هناك احتمال كبير أن لا يتحقق إفتـــراض تساوي التباين بين المجموعتين، لذلك يفضل في حالة عدم تساوي تباين المجموعتين قسمة الفـــرق الناتج عن اختلاف قيمة المتوسطات على الجذر التربيعي لما يسمى في حالة اختبار " ت " بالتباين الممزوج (pooled variance) والذي هو عباره عن دمج لتباين كلاّ المجموعتين المُمثلتين لمجتمعين مختلفين بتوزيعين مستقلين في توزيع واحد وبتباين واحد ويعبر عن هذا التباين وفق المعادله التاليه:

$$S_d^2 = \frac{S_1^2(n_1-1)+S_2^2(1-n_2)}{n_1+n_2-2}$$

حيث ان:

n_1 = هو عدد أفراد المجموعه الأولى

n_2 = هو عدد أفراد المجموعه الثانيه

S_1^2 = هو تباين المجموعه الأولى

S_2^2 =هو تباين المجموعه الثانيه

S_d^2 = هو التباين الممزوج

أما في حالة اختبار " ت " للعينات المترابطة فإنه عند استخدام مؤشر " d " لتقدير قيمة حجم الأثر فإنه يتم استخدام نفس القانون ولكن باستخدام المتوسطات القبليه والبعديه لنفس المجموعة(في حالة إجراء قياس قبلي وقياس بعدي لنفس المجموعة) لإستخراج الفرق وبالقسمه على الانحراف المعياري لاحد القياسيين (قبلي أو بعدي) في حالة الوفاء بإفتراض تجانس التباين، أو بالقسمه على الإنحراف المعياري للفروق في حالة استخدام الفروق بين القياس القبلي والقياس البعدي لإستخراج قيمة " ت " حيث تصبح معادلة مؤشر حجم الأثر " d " في هذه الحاله كما يلي:

$$d = \frac{\mu_f}{S_f}$$

حيث ان:

μ_f: متوسط الفروق بين القياس القبلي والقياس البعدي

S_f: الإنحراف المعياري للفروق

وهناك معادلة اخرى في حالة تساوي حجم المجموعة الضابطة والمجموعة التجريبية وهي:

(الضوى،٢٠٠٦)

$$ES = \frac{2r}{\sqrt{1 - r^2}}$$

حيث ان:

ES = حجم التأثير

r = معامل الارتباط بين درجات المجموعة الضابطة والمجموعة التجريبية.

وفي حالة عدم وجود انحرافات معيارية أو خطاء معيـــاري فيـــتم اســـتخدام المعادلـــة التالية(العبدالقادر،١٤٢٩):

$$d = t\sqrt{\left(\frac{n_E + n_c}{n_E n_c}\right)\left(\frac{n_E + n_c}{n_E + n_c - 2}\right)}$$

حيث أن:

d=قيمة حجم التأثير.

t=قيمة (ت) المحسوبة.

n_E=عدد المفحوصين في المجموعة التجريبية.

n_c= عدد المفحوصين في المجموعة الضابطة.

وهناك معادلة في حالة التصميم التجريبي لمجموعة واحدة(حالة العينة الواحدة) كما في حالة القياس القبلي pretest والقياس البعدي posttest (العبدالقادر،١٤٢٩):

$$ES = t\sqrt{\frac{2(1-r)}{n}}$$

حيث أن:

ES=حجم التأثير.

t=قيمة (ت) المحسوبة.

r=معامل الارتباط بين درجات القياسين القبلي والبعدي.

وأشار كوهين (١٩٨٨) إلى معايير للحكم على قيمة حجم الأثر المستخرجه بواسطة مؤشر" d " حيث اعتبره صغيراً عند القيمه ٠,٢٠ واعتبره متوسطاً عند القيمه ٠,٥٠ واعتبره كبيراً عند القيمه ٠,٨٠.

وهناك مؤشرات اخرى تستخدم لإستخراج حجم الأثر في حالة اختبار (ت) وهي:

- مؤشر جاما (y).
- مؤشر دلتا (Δ).

حيث أن مؤشر جاما (γ) لا يختلف عن مؤشر كوهين "d " والذي يعتمد على إستخراج الفرق بين المتوسطين ومن ثم قسمة الناتج على الإنحراف المعياري للتباين المزوج (Howell,1995)، وكذلك الحال بالنسبه لمؤشر دلتا (Δ) الذي يعتمد على نفس المبدأ (عوده والخليلي ،٢٠٠٠).

وأشار هويل (١٩٩٥) إلى أنه يمكن إعتماد نفس المعايير عند استخدام جاما (γ) كمؤشر لحجم الأثر ويمكن القول أن نفس المعايير يمكن أن تستخدم للحكم على مقدار حجم الأثر عند إستخدام مؤشر دلتا (Δ). ومن الجدير ذكره أن المعايير السابقه قد وُضعت على أساس دراسة قام بها كوهين حيث أثبت أنه عندما يكون حجم الأثر ٠,٢٠ فإن درجة التطابق (overlap) بين توزيع المجتمع الأول وتوزيع المجتمع الثاني (على أساس افتراض توزيعين لمجتمعين في حالة اختبار " ت "تصل إلى ٨٥% وأن درجة التطابق بين التوزيعين تصل إلى ٦٧% عندما تكون قيمة حجم الأثر ٠,٥ بينما تصبح درجة التطابق ٥٣% عندما تبلغ قيمة حجم الأثر ٠,٨.

٢- المؤشرات التي تُستخدم للدلاله على حجم الأثر في اختبار تحليل التباين الأحادي.

ميز الصياد (١٩٨٨م) عند استخدام اختبار تحليل التباين في حالة تساوي المتوسطات بين أربع حالات من حيث التصميم الإحصائي وهي :

أولاً : تحليل التباين أحادي العامل مع تساوي المفردات تحت كل مستوى من مستويات العامل (من المتغيرات المستقلة) .

ثانياً : تحليل التباين أحادي العامل مع عدم تساوي المفردات تحت كل مستوى من مستويات العامل المستقل .

ثالثاً : اختبار التأثير الرئيسي (Main Effect) في حالة التصميم العاملي (Factorial designe) .

رابعاً : اختبار التفاعل (Interaction) في التصميم العاملي .

تعتمد المؤشرات التي تُستخدم للدلاله على حجم الأثر في حالة استخدام تحليل التباين الأحادي لفحص فرضيات البحث على رصد نسبة التباين في المتغير التابع والمرتبط بتباين المتغير أو

المتغيرات المستقلة وتتراوح قيمة هذه المؤشرات من صفر إلى ١ . ومن الأمثله على هذه المؤشرات كما يشير كلاً من سنايدر ولاوسن (١٩٩٢)،و(نصار،٢٠٠٢):

- مؤشر مربع إيتا (Eta Squared) ويرمز له باللغة الإنجليزية بالرمز (η^2) و قدم هذا المؤشر العالم فيشرعام ١٩٣٢.
- ومؤشر مربع إيتا الجزئي (Partial Eta Squared) والذي قدمه كوهين عام ١٩٧٣ .
- ومؤشر مربع أوميجا (Omega Squared) والذي قدمه هيس عام ١٩٦٣.
- ومؤشر مربع إبسيلون (Epsilon Squared) والذي قدمه كيلي عام ١٩٣٥.
- ومؤشر مربع (ر) (R squared) والذي قدمه ستيفيتر عام ١٩٩٢.
- ومعادلة هيرزبيرغ (The Herzberg Formula) والتي قدمها هيرزبيرغ عام ١٩٦٩.
- معادلة لورد (The Lord Formula) والتي قدمها لورد عام ١٩٥٠.

وقد قدم الباحثون من خلال دراساتهم وأبحاثهم طرق مختلفة لحساب كل من المؤشرات السابقه ودللوا على أنه يمكن استخدام هذه المؤشرات في الأبحاث ذات التصاميم التجريبية والأبحاث الإرتباطية على حد سواء(سنايدر ولاوسن،١٩٩٢). هذا واضاف سنايدر ولاوسن أن بعض المؤشرات السابقة تضخم من نسبة التباين المفسر كما هو الحال مع مؤشر مربع إيتا ومؤشر (R squared)وهذا يعني أن قيم هذه المؤشرات والمستخرجه من خلال العينات هي قيم مضخمه، بحيث أنه إذا تم إستخراجها من خلال المجتمعات التي سحبت منها تلك العينات فإنها ستكون أقل من تلك التي تم تقديرها من خلال العينات.

ويشير كلاً من هويل (١٩٩٥م) ونصار(٢٠٠٢) الى أن أكثر المؤشرات شيوعاً لتقدير قيمة حجم الأثر في حالة استخدام تحليل التباين الأحادي هما مؤشر مربع إيتا ومؤشر مربع أوميجا وذلك لسهولة حسابهما.

وفيما يلي إستعراض لكلا المؤشرين:

أ- موشر مربع إيتا (η^2):

لايجاد قيمة مربع مربع إيتا (η^2) أو مايسمى بنسبة الارتباط من خلال تحليل التباين،فأننا نلجأ إلى المعادلة التالية:

$$Eta^2\ (\eta^2 = \frac{SS\ among}{\text{SS total}})$$

(المنيزل،٢٠٠٠)

حيث أن:

$SS\ among$= مجموع المربعات بين المجموعات.

SS tota = مجموع المربعات الكلي.

ويشير نصار (٢٠٠٢) الى أنه على الرغم من كون مؤشر مربع إيتا (η^2) متحيزاً(biased) حيث أنه يضخم من مقدار الإرتباط بين تباين المتغير التابع و تباين المتغير المستقل والمقدر من خلال العينه مقارنة بالقيمه التي يتم إستخراجها من خلال جميع أفراد مجتمع الدراسة ،وتعتبر عملية حساب هذا المؤشر في غاية السهوله ومن السهل تفسير القيمه التي يتم الحصول عليها من خلالها. يتم الإستدلال عند إستخدام أسلوب تحليل التباين الأحادي على مقدار التباين الكلي في بيانات الدراسة من خلال ما يعرف بمجموع مربع الإنحرافات الكل(Total sum of squares) ويعود جزء من هذا التباين الكلي في البيانات إلى أثر المعالجة حيث يفسر الإختلاف في درجات مجموعات الدراسة المختلفة اى أثر هذه المعالجة ويمكن أن يعود جزء آخر من التباين الكلي إلى متغيرات مستقله أو ثانويه أخرى في الدراسة وإلى التفاعل ما بين هذه المتغيرات بينما يعود باقي التباين الى ما يسمى بالخطأ العشوائي. ولنفرض أنه في دراسة ما كان هناك متغير مستقل واحد (معالجه) ومتغير تابع واحد ففي هذه الحاله يقسم التباين الكلي إلى تباين ناتج عن أثر المعالجه أو تباين مفسر بينما يعود باقي التباين إلى ما يسمى بالخطأ. وبناءً عليه تمت صياغة معادلة مربع إيتا السابقة.

أما في حالة القياسات المتكررة يتم حساب مربع ايتا من خلال الصيغة التالية :

$$(\eta^2) = \frac{SS_A}{SS_A + SS_{AXP}}$$

(الكناني،٢٠٠٢)

SS_A = مجموع المربعات بين المجموعات.

SS_{AXP}= مجموع المربعات التفاعل المستقل A و الأفراد.

مميزات مربع إيتا (η^2):

ذكرت الصائغ(١٤١٧ه) مجموعة من المميزات لمربع ايتا منها:

١- يُعتبر مقياساً مقبولاً وخصوصاً في البحوث التربوية لحساب الدلالة العملية أو حجم التأثير والمسمى بالتباين المفسر، ويكون حسابها بدون صعوبة من جدول ANOVA الكامل في حالة اختبار تحليل التباين الأحادي.

٢- أن معظم الإحصائيين يوافقون على أن مربع ايتا يعطي معلومات أكثر دقة من قيمة الفرق الإحصائي حيث تأخذ في الاعتبار حجم العينة .

٣- تظهر أهميته وفائدته عند مراجعة الدراسات السابقة باستخدام ما يسمى بأسلوب ما وراء التحليل أو التحليل البعدى وهي طريقة كمية لمراجعة الدراسات السابقة في أي مجال من المجالات بطريقة تكاملية ،وبأقل قدر من التحيز سواء كان هذا التحيز مراجعة ذاتية الباحث الذي يقوم بالمراجعة ،أو تحيز يرجع للدراسات السابقة موضع المراجعة .

٤- يبقى لقيمة مربع إيتا دور كبير في تقييم الدلالة العملية و في بيان درجة التنبؤ الفعلي للدراسة حتى لو كانت قيمتها منخفضة للغاية .

٥- أن قيمة مربع إيتا تساوي قيمة معامل التحديد R^2 المستخرج من نموذج تحليل الانحدار الذي يتم استخدامه عند تحليل الاختلاف في تصميم يشبه تحليل ANOVA .

٦- يشكل مربع إيتا رابطة قوية بين التحليل التقليدي للتباين من ناحية وبين الانحدار المتعدد من ناحية أخرى .

٧- يستخدم مربع إيتا في تفسير أهمية البحث وفي قياس الفروق التي يظهرها التحليل الإحصائي وقد أصبح استخدامه شائعاً خاصة في مجال العلوم الاجتماعية نظراً للتشابه مع معامل التحديد R^2 الأكثر فائدة والناتج من تحليل الانحدار .

ب- مؤشر مربع أوميجا (ω^2):

يذكر الكيلاني والشريف (٢٠٠٧)عن مربع أوميجا (ω^2) أنه النسبة من التباين الكلـــي التي يمكن تفسيرها في المتغير التابع في المجتمع الاصلي الذي اختيرت منه العينة،ويعتـــبر بـــارمتراً وينتمي إلى الإحصاء الإستدلالى أي عكس مربع إيتا الذي ينتمي بدوره إلى الوصـــفي. ويـــشير نصار(٢٠٠٢) الى ان مؤشر مربع أوميجا يعتبر أقل تحيزاً من مؤشر مربع إيتا.

ولحساب قيمة مربع أوميجا يتم إستخدام المعادلة :

(الكيلاني والشريف ،٢٠٠٧)

$$\omega^2 = \frac{V_b - (F-1)}{(V_b.F) + V_w + 1}$$

حيث أن:

V_b = درجات الحرية للبسط (بين المجموعات).

V_w = درجات الحرية للمقام (داخل المجموعات).

F = قيمة (ف) لتحليل التباين.

وبخصوص الحكم على مقدار مؤشر مربع إيتا فقد أشار كوهين (١٩٧٧)إلى إعتبار هـــذا المقدار صغيراً عندما يساوي (٠,٠١) وإلى إعتباره متوسطاً عندما يساوي (٠,٠٦) وإلى إعتباره كبيراً عندما يساوي (٠,١٤).

أما بالنسبه للحكم على مقدار مربع أوميجا فقـــد ذكـــر (نـــصار،٢٠٠٢) في بحثـــه أن الدراسات لم تشر إلى معايير محدده للحكم على مقدار مربع أوميجا ولعل ذلك مرده إلى أن مربع إيتا أكثر إستخداماً من مؤشر أوميجا وربما يعود ذلك لأن عملية حسابه أكثر ســـهولة، لكـــن ونتيجةً للتشابه بين مؤشر مربع إيتا ومؤشر مربع أوميجا من حيث طريقة التفسير ونتيجةً لتقارب قيمة المؤشرين وخاصةً مع العينات الكبيره (التي تزيد عن ١٠٠) فإنه يمكن إستخدام نفس المعايير التي قدمها كوهين في حالة مربع إيتا للحكم علـــى مقـــدار قيمـــة مربـــع أوميجـــا(ســـنايدر ولاوسن،١٩٩٢).

المبحث الثاني

مقدمة:

لقد قام الباحث في هذا المبحث باستعراض بعض الأساليب الإحصائية الشائعة الاستخدام وبالتحديد اختبار " ت " (T-test) واختبار تحليل التباين(ANOVA).

أولاً: اختبار (ت) T-test:

يعد هذا الاختبار من أكثر الأساليب الإحصائية شيوعاً واستخداماً في مجال البحوث التجريبية التربوية والنفسية،حيث أشار كلا من فرج (١٩٩٦م)،و(الأعسر، ٢٠٠٢م) وآخرون إن الفضل يعود في اشتقاق هذا الاختبار إلى العالم الايرلندي وليلم كوسيت (W.S Gosset) عام ١٩٠٨م، حيث نشر بحثا باسم مستعار هو Student وعُرف هذا التوزيع بالاسم ((Students T-Distibution)) ويختصر باسم توزيع (T)، والذي توصل من خلاله إلى اشتقاق معادلة للتوزيع الاحتمالي (ت)حيث قيمته هي:

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

افتراضات استخدام اختبار (ت) T-test :

يشير عوده (٢٠٠٠م) إلى أن هذا الاختبار يستند إلى توفر عدد من الافتراضات و هي:

أ – مستوى القياس :

لاستخدام هذا الاختبار يشترط أن تكون البيانات فئوية (فتريه) أو نسبية.

ب – أن يكون حجم العينة صغيراً :

وهنا يتطلب هذا الافتراض إن يكون حجم العينة (n) اقل من ٣٠ واكبر من ٥، وحيث أن حجم (n) لا يسمح بتقدير جيد ل (σ) لذلك لتجنب الخطأ نستعيض عن(σ) بالمعلمة (s).

ويفضل في أي دراسة تستخدم اختبار (T) أن يكون الفرق بين حجمي عينتي البحث صغيراً لأن لحجم العينة تأثيراً على مستوى الدلالة (α) وعلى كل من المتوسط والتباين (السيد، ٢٠٠٥) .

ج - التوزيع الطبيعي:

ويقتضي هذا الافتراض أن عدد المشاهدات (س١)في المجتمع الأول تتخذ شكل التوزيع الطبيعي لوسط يساوي (م١) وكذلك الأمر بالنسبة للمشاهدات (س٢) في المجتمع الثاني يفترض فيها أن تتخذ شكل التوزيع الطبيعي لوسط يساوي (م٢).

يعتمد توزيع (T) على افتراض أن العينة المسحوبة من مجتمع إحصائي توزيعه معتدل أو قريب من الاعتدال. بناءً على ذلك يستطيع الباحث استخدام توزيع (T) لإجراء اختبارات الفروض وإيجاد حدود الثقة لمتوسط المجتمع أو الفرق بين متوسطين عندما تكون أحجام العينات صغيرة وتباينات المجتمعات مجهولة. أما إذا كان توزيع المجتمع الأصلي الذي سحبت منه العينة بعيداً عن الاعتدال بدرجة كبيرة فمن الأفضل استخدام الاختبارات اللامعلمية لإجراء اختبارات الفروض (تشاو، ١٩٩٠م) .

د - تجانس التباين في المجتمعين :

يتوجب هذا الافتراض أن يكون تباين المشاهدات في العينة الأولى S_1^2 لا يختلف عن تباين المشاهدات في العينة الثانية S_2^2 وفي هذه الحالة يقال أن العينتين متجانستان.

اقترح هارتلي (Hartly,1940) طريقة لاختبار التجانس وهـي حساب قيمة (ف) مـن قسمة التباين الأكبر على التباين الأصغر من تباينات المجموعـتان ثم مقارنة النـاتج بتوزيع خاص يسمى F-Max بدرجات الحرية (ك (ن-١)) حيث " ك " هذه هي عدد المجموعـات ، و"ن" حجم العينة ، وهذا الاختبار يكفي للتعرف على مدى التجانس (مراد، ٢٠٠٠م).

ويذكر كـلا مـن (الـشربيني،٢٠٠٧م)،(عـودة والخليلـي ،٢٠٠٠) ،(ابوحطـب ١٩٩١،م)،(علام ،٢٠١٠)... وآخرون أن افتراض تجانس التباين يمكن التغاضـي عنـه حـين يتساوى حجما العينتين أي ($n_1 = n_2$).

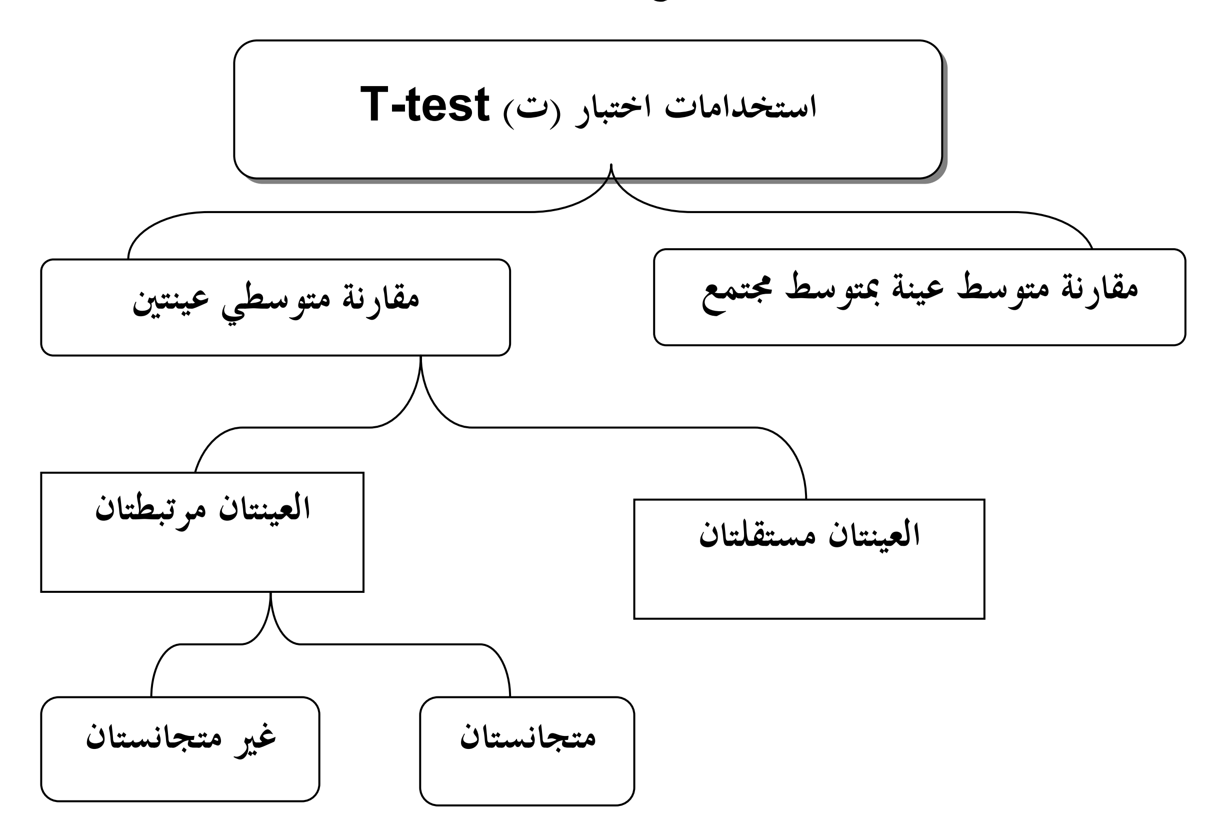
هـ - الاستقلالية :

الاستقلالية لا تعني استقلالية البيانات بين المجتمعين فقط ، بل تعني استقلالية المشاهدات ضمن المجتمع الواحد أيضاً كعملية تطبيق اختبار قبلي و اختبار بعدي على مجموعة واحدة.

يشير أبوحطب (١٩٩١م) الى أن الاستقلال يعني ببساطة أن البيانات التي نجمعها سواءً بين المجموعات أو داخل المجموعات ليست متزاوجة أو متكررة أو متداخلة أو معتمدة على بعضها البعض على أي نحو.ولا يتوفر ذلك إلا إذا كان اختبار العينات عشوائيا تماما، أي تحكمه عوامل المصادفة من ناحية، وكذلك أن يكون قد استخدم وسائل الضبط التجريبي من ناحية أخرى .فإذا تزاوجت الدرجات على نحو أو آخر ، سواء أكان ذلك عن طريق تكافؤ المجموعات أو تكرار الملاحظات على نفس الأفراد فإن المجموعات حينئذً تكون مرتبطة وفي هذه الحالة لابد من استخدام اختبار (ت) للمجموعات المرتبطة أو للقياسات المتكررة.

أنواع استخدامات اختبار (ت) T-test

شكل (١) أنواع استخدامات اختبار (ت)



أ- مقارنة متوسط عينة بمتوسط مجتمع:

عند مقارنة متوسط عينة عشوائية مع متوسط المجتمع الأصلي الذي سحبت منه هذه العينة يجب التمييز بين حالتين :

الحالة الأولى : إذا كان تباين المجتمع معلوما وحجم العينة صغير ($n \le 30$)يستخدم اختبار Z-test .

الحالة الثانية : أن يكون تباين المجتمع مجهولا وحجم العينة صغير ($n \le 30$) وهنا يستخدم اختبار T-test.

مقارنة متوسط عينة بمتوسط مجتمع غير معلوم تباينه: ومثال على ذلك نفرض أن لدينا مجتمع طلاب التربية الفكرية بالقنفذة متوسط الذكاء فيه μ وتباينه σ^2 غير معلوم ، وأراد احد الباحثين تقنين اختبار للذكاء على نفس هذا المجتمع ،فاختار عينة عشوائية من المجتمع فكان متوسط الذكاء فيها هو $\overline{X}$ وتباينه S^2، واراد مقارنة متوسط ذكاء العينة مع متوسط ذكاء المجتمع، فهنا نفتقد إلى الانحراف المعياري للمجتمع،وعليه يجب علينا استخدام T-testوذلك لمقارنة متوسط ذكاء العينة مع متوسط ذكاء المجتمع المسحوبة منه العينة .

ولكن قبل البدء في استخدام T-test على بيانات البحث يجب التأكد من استيفاء بيانات العينة لافتراضات الأعتدالية المذكورة سابقاً، وبعد توفرها يمكن للباحث أن يستكمل إجراء T-test.وفي هذا المثال يكون الهدف التحقق من صحة الفرض الصفري:

متوسط ذكاء العينة لا يختلف عن متوسط ذكاء المجتمع $H_o: \mu = \mu_o$

يستخدم قانون حساب قيمة T-test :

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

(الشربيني ،٢٠٠٧م)

حيث أن T:اختبار دلالة الفرق بين متوسط العينة$\bar{X}$ ومتوسط المجتمع μ.

$\bar{X}$:متوسط العينة.

μ:متوسط المجتمع.

S:الانحراف المعياري للعينة (الانحراف المعياري=الجذر التربيعي للتباين).

n: عدد أفراد العينة .

نقارن قيمة T المحسوبة من القانون السابق بقيم t من الجدول أمام درجات حرية (١-n)=V تحت مستوى الدلالة المطلوب.

يرفض الفرض الصفري إذا كانت Tالمحسوبة $\leq$ قيمة t من الجدول ، حيث تقع T المحسوبة في هذه الحالة داخل منطقة الرفض ، ويقبل الفرض البديل .

أما إذا كانت قيمة T المحسوبة <قيمة t الجدولية فتكون داخل منطقة القبول ، ونقبل الفرض الصفري ونرفض الفرض البديل .

ب – مقارنة متوسطي عينتين :

وهنا يجب التمييز بين نوعين من العينات وهما:

أولاً: الفرق بين متوسطي عينتين مستقلتين : Independent Samples

في اغلب الدراسات التربوية نحتاج إلى مقارنة متوسطي عينتين مستقلتين كمقارنة آراء المعلمين وآراء المشرفين التربويين حول الأساليب الإشرافية .

في المثال السابق نجد أن بيانات كل من العينتين المعلمين و المشرفين. الضابطة والتجريبية تحسب باستقلالية عن الأخرى. أي أنه لا توجد أي علاقة ارتباط بين مشاهدات العينة الأولى ومشاهدات العينة الثانية. لذا ينبغي في هذه الحالة استخدام اختبار (ت) لعينتين مستقلتين ويمكن توضيح لقانون حساب قيمتي (ت) وطريقة إجراء الاختبار:

أ- العينتان مستقلتان ومتجانستان:

إذا كان نتيجة اختبار (لفن) لتجانس التباين غير دالة إحصائياً بالتالي تستخدم قانون اختبار (ت) لعينتين مستقلتين ومتجانستين. وفي هذا الأسلوب يتم مزج تباين العينتين أو خلطهما معاً بحيث يعطي تقديراً أفضل وأكثر كفاءة لتباين العينتين ويتم ذلك من خلال دمج (إضافة) مجموع المربعات في البيانات من كل من العينتين والقسمة على درجات الحرية الإجمالية لكل من العينتين أي من خلال العلاقة التالية:

(الشربيني ، ٢٠٠٧م)

$$T = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{S_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

حيث S_P هو التباين المشترك أو الممزوج .

ويمكن الحصول على التباين المشترك من القانون التالي:

(تشاو،١٩٩٠)

$$S_P^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

ثم تقارن قيمة t المحسوبة من القانون مع القيمة الحرجة $t_{(n1+n2-2,\alpha)}$ فإذا جاءت $T<t$ فيمكن القول أن الفرق بين متوسطي العينتين غير دال إحصائياً . أما إذا جاءت $T \geq t$ يمكن القول أن الفرق بين متوسطي العينتين دال إحصائياً .

ب- العينتان مستقلتان وغير متجانستين :

هنا تكون نتيجة اختبار لفن لتجانس التباين دالة إحصائيا أي أن العينتين مستقلتان وغير متجانستين .فنحسب قيمة T بدون مزج تباين العينتين باستخدام القانون التالي

(عوده والخليلي،٢٠٠٠)

$$T = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{S_P\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

حيث : $\bar{X}_1, \bar{X}_2$: متوسط العينة الأولى والثانية على التوالي .

و S_1^2, S_2^2 تباين العينة الأولى والثانية على التوالي .

n_1 , n_2 : حجم العينة الأولى والثانية على التوالي.

أما القيمة الجدولية فنحصل عليها بدرجات الحرية γ التي تحسب بالقانون التالي:

$$\gamma = \frac{\left[\left(S_1^2 / n_1\right) + \left(S_2^2 / n_2\right)\right]^2}{\frac{\left(S_1^2 / n_1\right)^2}{n_1 + 1} + \frac{\left(S_2^2 / n_2\right)^2}{n_2 + 1}} - 2$$

(عوده والخليلي،٢٠٠٠)

حيث :

γ: متوسط العينة الأولى والثانية على التوالي .

و $S_1^2 : S_2^2$ تباين العينة الأولى والثانية على التوالي .

n_1 , n_2 :حجم العينة الأولى والثانية على التوالي.

وهذه الطريقة التقريبية لحساب درجات الحرية للعينتين المستقلتين .

تستخدم قيمة درجة الحرية y الممزوجة لإيجاد قيمة t الجدولية عند مستوى الدلالة المطلوب α.

فإذا جاءت $T<t$ فيمكن القول أن الفرق بين متوسطي العينتين غير دال إحصائياً . أما إذا جاءت $T \geq t$ فإن ذلك دلالة على أن الفرق بين متوسطي العينتين دال إحصائياً.

ثانياً: العينتان المرتبطتان (غير المستقلتين) Correlated Samples .

سبق وان ذكرنا فيما سبق ضرورة مراعاة فرضيات استخدام اختبار t ومن هذه الفرضيات الاستقلالية أي ضرورة استقلال بيانات العينتين ، ولكن من الناحية التطبيقية يصادف الباحث كثير من التجارب أو الدراسات تكون العينتان مترابطتين (غير مستقلتين) إذا وجد

ارتباط فيما بين بياناتهما وينشأ هذا الارتباط عندما يتم تكرار الاختبار على المجموعة نفسها مرتين كاختبار قبلي واختبار بعدي. أو عند أخذ التوائم ووضع كل منهما في مجموعة. أو عند أخذ أزواج متطابقة ووضع كل فرد من الأزواج في إحدى المجموعتين .

وبصورة إجمالية تكون البيانات غير مستقلة إذا كان بالإمكان أن نربط بين مشاهدات المجموعة الأولى ومشاهدات المجموعة الثانية. و في مثل هذه الحالات تكون ($n_1 = n_2 = n$) ويكون لمعامل الارتباط بين المشاهدات في المجموعتين قيمة تختلف عن الصفر (عودة والخليلي ، ٢٠٠٠م).

وقبل البدء في استخدام اختبار (T) لعينتين غير مستقلتين مترابطتين ينبغي التأكد من استيفاء بيانات العينتين لافتراضات الاختبار (T) وهي الاعتدالية، وتجانس التباين، بالإضافة إلى الارتباط بين مشاهدات العينتين . بينما تجدر الإشارة الى إن تساوي حجمي العينتين المترابطتين يساعد على التنازل عن افتراض تجانس التباين.

ثانيا :تحليل التباين :analysis of variance

يذكر علام (٢٠١٠م) أن فيشر Fisher عالم الإحصاء الانجليزي قد صمم هذا الأسلوب لتيسير تحليل البيانات المستمدة من التجارب الميدانية والمختبرية في مجال البحوث الزراعية والبيولوجية ،وتفسير نتائجها. ويُمثل اليوم أحد أدوات البحث المهمة ليس فقط في هذه المجالات ، وإنما في مختلف مجالات العلوم النفسية والتربوية والاجتماعية والطبيعية وغيرها.

كما يذكر أبو حطب وآخرون(١٩٩١م) بأن تحليل التباين " يعد أسلوباً إحصائياً لازماً لفهم طبيعة المنهج التجريبي وشبه التجريبي."

في حين أشار عدس(١٩٩٧م) إلى أن تحليل التباين " يقدم لنا أسلوبا إحصائياً مناسباً لمقارنة عدة متوسطات مع بعضها البعض في نفس الوقت."

وكذلك يعرف عوده و الخليلي (٢٠٠٠م) "تحليل التباين مبدئياً طريقة ذكية لاختبار اختلاف أوساط مجموعتين أو أكثر دفعة واحدة من خلال التباين."

أما طه والقاضي (١٩٩٤ م)فيعرفانه بأنه " أسلوب إحصائي الهدف منه تقسيم مجموع مربعات الانحرافات الكلي إلى مكوناته الأساسية ومن ثم إرجاع كل هذه المكونات إلى سببه."

والغرض من تحليل التباين يتمثل في دراسة وتحليل أثر متغير أو أكثر من المتغيرات الوصفية على متغير كمي ، حيث يستخدم هذا الأسلوب لمعرفة هل هناك فروق بين متوسطات المجموعات أم لا توجد .

و طريقة تحليل التباين تتمثل في حساب المجموع الكلي لمربعات الانحرافات لجميع الوحدات التجريبية في التجربة عن المتوسط العام ومن ثم تقسيمه إلى مكوناته طبقاً للمصادر المسببة لها والتي يختلف عددها من تجربة لأخرى بحسب ظروف ونوع وتصميم التجربة كذلك يتم بنفس الطريقة تقسيم درجات الحرية الكلية ثم بعد ذلك تدون النتائج في جدول يطلق علية اسم جدول تحليل التباين ANOVA (طه والقاضي،١٩٩٤ م).

وتعتمد الطريقة الإحصائية لتحليل التباين على حساب تباين الخطأ وذلك بحساب المربعات داخل المجموعات، وحساب التباين المفسر وذلك بحساب المربعات بين المجموعات، وحساب درجات الحرية لتحويل تلك المربعات إلى التباين المقابل لها، وللكشف عن الدلالة الإحصائية للنسبة الفائية ،حساب النسبة الفائية، والكشف عن دلالتها الإحصائية(السيد، ٢٠٠٥ م) .

حيث بين علام (٢٠١٠م) بأن النسبة الفائية الملاحظة =

$$\frac{\text{تقدير تباين المجتمع باستخدام العينة الأولى}}{\text{تقدير تباين المجتمع باستخدام العينة الثانية}}$$

$$\text{والنسبة الفائية} = \frac{\text{متوسط مجموع المربعات بين المجموعات}}{\text{متوسط مجموع المربعات بين المجموعات}}$$

فوائد تحليل التباين في العلوم التربوية:

توجد الكثير من الفوائد الإحصائية لتحليل التباين ومنها:

١- يُمكن استخدامه في قياس الفروق الفردية في السمات الشخصية والعقلية وكذلك في اختلاف الأداء وقياس الدلالة الإحصائية لذلك.

٢- يُفيد في قياس مدى تجانس عينات الباحثين وكذلك معرفة تجانس المفردات التي تتكون منها الاختبارات.

٣- يُستخدم في قياس الفروق الفردية والجماعية نظراً لأنه يعتمد على استخدام حساب مدى انحراف كل فرد عن متوسط الأفراد الآخرين أو انحراف كل جماعة عن متوسط الجماعات الأخرى.

٤- يفيد في قياس عوامل الخطأ للفروق الناتجة من اختلاف المجتمعات الأصلية للعينات(الزراد، ١٩٨٨ م).

أسباب استخدام تحليل التباين بدلاً من استخدام اختبار (ت) :

يفضل استخدام تحليل التباين(F) بدلاً من اختبار (T) وذلك للأسباب التالية:

- الجهد المبذول في عمل المقارنات،عدد المقارنات

$$= \frac{n(n-1)}{2}$$

حيث انه كلما زاد عدد العينات زاد عدد مرات إجراء اختبار النسبة زيادة ملحوظة.

- ضعف عملية المقارنة،في اختبار(T) يتم المقارنة بين كل متوسطين لمجموعتين على حده وبالتالي تهمل بقية المعلومات عن المجموعات الأخرى مؤقتاً والتي من الواجب أخذها بعين الاعتبار لأنها جزء يجب ألا ينفصل وبالتالي فهي تؤثر على قوة المقارنة.
- مخاطرة الوقوع في خطأ من النوع الأول، ذكر الشربيني (٢٠٠٧ م) بأن " عدد المقارنات ومستوى الدلالة يرتبطان باحتمالية الوقوع أو ارتكاب خطأ أو أكثر من النوع الأول طبقاً للعلاقة التالية :

احتمالية الوقوع في خطأ من النوع الأول .

$$= n \ (\alpha - 1)$$

حيث أن :

n = عدد المقارنات.

α = مستوى الدلالة المستخدم في هذه المقارنات."

متطلبات تحليل التبـاين :

يذكر علام (٢٠١٠م) إن تحليل التباين يستند إلى بعض الفروض التي لا تختلف كثيراً عـــن فروض النسبة التائية في حالة عينتين .

فيما يشير فرج (١٩٩٦م) إلى أنه حتى يمكننا استخدام تحليل التباين وفق أصول منهجيـــة سليمة يتعين الالتزام بعدد من المتطلبات الأساسية فيه وأهمها التالي:

١) عشوائية سحب المجموعات من مجتمع اعتدالي:

بحيث يجب أن نختار أفراد المجموعات المختلفة على أسس عشوائية

٢) تجانس تباين العينات :

بالإضافة إلى ذلك يجب أن يكون تباين المجموعات متجانساً.

٣) استقلال تباين المجموعات :

وهذا من الشروط الجوهرية في تحليل التباين، وتتضح أهمية هذا الشرط في الفرض الأساسي لهذا الاستقلال في أن عدم الالتزام به ، يترتب عليه عدم مطابقة النسبة بين تباين ((بين المجموعات)) إلى تباين ((داخل المجموعات))،لتوزيع (F) وبالتالي تصبح دلالة النتائج مشكوك فيها.

اختبـار افتراضات استخدام تحليل التبـاين :

١ - التحقق من تجانس التباين :

أ- اختبار هارتلي Hartley:(في حالة عدد أفراد العينات >< ٣٠)

يذكر علام (٢٠١٠) انه يستخدم لتحقق من تجانس التباين لعينات متعددة،كما يمكن أن يستخدم في حالة عينتين ويسمى ((اختبار النسبة الفائية القصوى)).

ب- اختبار كوكران Cochran:(في حالة عدد أفراد العينات >< ٣٠)

ويستخدم في حالة العينات المتساوية في الحجم ولها توزيع ملتوي أو مفرطح. وعلى الرغم من أنه يفضل استخدام كوكران في حالة العينات متساوية الحجم ألا انه يصلح لعينات مختلفة الحجم (علام ،٢٠١٠).

وكذلك يستخدم في حالة العينات غير المتساوية في الحجم وتتبع التوزيع الطبيعي.

ج- اختبار شيفية Scheffe:(في حالة عدد أفراد العينات >< ٣٠) .

يستخدم في حالة العينات غير المتساوية في الحجم ولا تتبع التوزيع الطبيعي.

٢ - التحقق من استقلالية العينات:

يتم التحقق من استقلالية العينات بعضها عن بعض عن طريق اختبار مربع كاى Chi-Sguare).

٣- التحقق من كون العينات تتبع التوزيع الطبيعي:

أ - معامل الالتواء: ويتم استخدام معامل الالتواء إذا كان عدد أفراد العينات اكبر من أو يساوي ٣٠.

ب- اختبار مربع كاى (Chi-Sguare):ويتم استخدام مربع كاى إذا كان عدد أفراد العينات أصغر من ٣٠ وأكبر من أو يساوي ٢٠.

ج- اختبار سمير نوف-كلموجورف(k.s): يمكن استخدام اختبار سمير نوف-كلموجورف إذا كان عدد العينات أقل من ٢٠.

يذكر الزرآد(١٩٨٨) بأنه " يمكن حساب مدى مطابقة التوزيع التكراري المشاهد للتوزيع الاعتدالي باستخدام اختبار سمير نوف(Smirnov) وكلموجورف(Kolmogorov) والذي يرمز له باختبار(k.s)".

البدائل الممكنة في حالة مخالفة افتراض أو أكثر من افتراضات تحليل التباين :

نظراً لان الإخلال بأحد هذه الافتراضات وعدم تحقيق البيانات لهذه الافتراضات قد يؤدي إلى التشكيك في مصداقية النتائج عند استخدام تحليل التباين وبالتالي فيمكن استخدام بدائل ممكنة على النحو التالي:

١- في حالة أذا كانت البيانات غير فئوية أو نسبية:

يستخدم الإحصاء اللامعلمي كروسكال واليس ،حيث يذكر(علام، ٢٠٠٠ م)بأنه " يعد اختبار كروسكال واليس بديلاً لا معلمياً لتحليل التباين أحادي الاتجاه."

٢-في حالة إذا كانت البيانات لا تتبع التوزيع الطبيعي:

يتم استخدام التحويلات وذلك من اجل الاقتراب من اعتدالية التوزيع أو استخدام الإحصاء اللامعلمي الذي لا يشترط إعتدالية التوزيع وهناك عدة طرق لتحويل البيانات أهمها:

أ- تحويله الجذر ألتربيعي Square Root Transformation

ب- التحويلة اللوغاريتمية Logarithmic Transformation

ج- تحويلة المقلوب Reciprocal Transformation

د- تحويلة الدالة العكسية لجيب الزاوية Angular or Inverse Sin Transformation

٣- في حالة إذا كانت البيانات لا تحقق افتراض تجانس التباين:

ويمكن استخدام عدة طرق منها:

أ- استخدام التحويلات المختلفة.

ب- استخدام الإحصاء اللامعلمي (كروسكال واليس).

ج- استخدام طريقة البوتستراب.

وبعد التأكد من تحقق الافتراضات السابقة في البيانات وبعد تصحيحها في حالة المخالفة يمكن بعد ذلك استخدام تحليل التباين ومن ثم إعطاء نتائج صحيحة.

إن أساليب تحليل التباين تتميز بالمرونة بحيث يمكن استخدامها في تصميمات تجريبية متعددة، مثل : تصميم العامل الواحد single-factor Design ، والتصميمات ألعامليه factorial Designs ، وتصميمات القياسات المتكررة (علام ، ٢٠١٠).

تحليل التباين أحادي الاتجاه One Way ANOVA: :

أن تحليل التباين لاتجاه واحد يفيد الباحث في " اختبار الاختلافات بين عدد من المجموعات في متغير تابع واحد عوده،والخليلي (٢٠٠٠ م). حيث ان هناك أشكالاً مختلفة لتحليل التباين تتوقف هذه الأشكال على عدد المتغيرات المستقلة والتابعة وأحد هذه الأشكال هو تحليل التباين في اتجاه واحد والذي يهتم بالكشف عن الفروق أو الاختلافات بين عدد من المجموعات في متغير تابع واحد وكل مجموعة من هذه المجموعات يطلق عليها معالجة (الشر بيني، ٢٠٠٧ م).

وبناءً عليه يتبين أن تحليل التباين أحادي الاتجاه هو احد أنواع تحليل التباين يركز على دراسة تأثير متغير مستقل واحد له عدة مستويات على متغير تابع واحد.

حالات تحليل التباين الأحادي:

الحالة الأولى: تحليل التباين الأحادي في حالة تساوي حجوم العينات.

إذا كان لدينا عينات عشوائية متساوية الحجوم وكان حجم كل منها n ،وكانت متوسطاتها (μ1, μ2 , μ3 , ………… , μk) وتباين σ^2 حيث:

$$\sigma_1^2 = \sigma^2 = \sigma_2^2 \ldots\ldots \sigma_k^2.$$

لاختبار الافتراضية : Ho: μ1 = μ2 = μ3 = ………… = μk

حيث أن :Kعدد المجموعات،nعدد أفراد المجموعة الواحدة،Xij يرمز لكل عنصر

i =1 , 2 , …… , k

j =1 , 2 , …… , n

وبالتالي:

جدول (٢)

تحليل التباين الأحادي في حالة عدم تساوي حجوم العينات

العناصر n	المجموعات Groups K
	1 2 3 4 .... I... .... k
1	X11 X21 .... Xi1 ....XK1
2	X12 X22.............. Xi2 .... Xk2
3	X13
.	
J	X1j
.	
N	X1n
∑xi	∑x1 ....∑xi ... .∑xk
µx	X1 .... Xi ....Xk

الاختلاف بين قيم Xij يعزى إلى:

١- الاختلاف بين قيم Xij الواقعة ضمن المجموعة الوحدة(تباين الخطأ)

٢- الاختلاف بين المجاميع نفسها. (الشمراني،١٤٢١هـ)

وبذلك يصبح جدول تحليل التباين لاتجاه واحد كما يلي:

جدول (٣)

تحليل التباين في اتجاه واحد

F tab	F cal	متوسط المربعات	مجموع المربعات	درجة الحرية	مصدر الاختلاف
$F_{(\mu,v1,v2)}$	$\frac{MS_B}{MS_w}$	MS_B	SS_B	k- 1	بين المجموعة
————	————	MS_W	SS_W	N – K	ضمن المجموعة
————	————	————	SS_T	N –1	الكلي

ويكون القرار هو رفض H_0 إذا كانت قيمة F المحتسبة أكبر من القيمة الجدولية (1– K(n
F α , k- 1 ,

الحالة الثانية : في حالة عدم تساوي حجوم العينات: يتبع نفس الأسلوب عند تساوي حجـــوم العينات ، ولا كن يصبح حجم العينةni بدلاً مـــن n .

المقارنات البعدية : Posteriori comparisons

بعد استخدام أسلوب تحليل التباين لاختبار تساوى عدة متوسطات وفي حالة رفض الفرض الصفري يأتي دور إجراء المقارنات البعدية والتي تهدف إلى إجراء المقارنات الفردية بين المتوسطات لتحديد اتجاه الفرق بين المجموعات وهناك حالتين للمقارنات هي :

أولاً :طرق المقارنات المستخدمة مع الاختبارات المعلمية:

١- طريقة أقل فرق دال معنوي (L.S.D).

٢- طريقة توكي للفرق الدال الصادق (H. S. D) .

٣- طريقة شيفية Scheffes Method.

٤- طريقة نيومان – كولز Newman – Kuls.

٥- طريقة دنكن Duncans Method.

ثانياً :المقارنات المستخدمة في حالة الاختبارات اللامعلمية:

١- اختبار مان وتني .

٢- اختبار نيومان Newman .

الدراسات السابقة

مقدمة :

هناك العديد من الدراسات قد تم تناولها من خلال الفترة الزمنية من الأقدم إلى الأحدث.

نشر تشيس وتشيس (Chase&Chase,1976) دراسة هدفت الى معرفة قوة الاختبارات الإحصائية المستخدمة في البحوث المنشورة في مجلة علم النفس التطبيقي لسنة ١٩٧٤م وتكونت عينة الدراسة من الاختبارات الإحصائية وعددها ٣٣٧٣ اختباراَ وهي الاختبارات المستخدمة في ١٢١ مقالة من أصل ١٦٢ ، وأشارت نتائج الدراسة الى أن الاختبارات ذات الدلالة الإحصائية لها دلالة عملية متوسطة ،وعالية .كما أن قوة الاختبارات كانت عالية .

واجرى هاس وآخرون (Haase and waechter and Solonon,1982) دراسة بعنوان واقع الدلالة العملية بواسطة مربع آيتا للاختبار الإحصائي (F) المستخدمة في مجلة علم النفس الإرشادي للمجلدات (١٧-٢٦) وللفترة الواقعة مابين ١٩٧٠-١٩٧٩ حيث تكونت عينة الدراسة من ١٠٧ مقال بواقع ١٠٤٤ اختباراَ إحصائياً. وتوصلت الدراسة الى أن حجم التأثير محسوباً بمربع آيتا للإحصائي (F) كان صغيراً ٠،٠٨ .

وفي دراسة حجيمات وعليان (١٩٩٧م) بعنوان (دراسة واقع الدلالة الإحصائية العملية وقوة الاختبار الإحصائية المستخدمة في رسائل ماجستير الإرشاد النفسي والتربوي بالجامعة الأردنية) والتي هدفت الى تناول واقع الدلالة الإحصائية وعلاقتها بكل من الدلالة العملية وقوة الاختبار وتكونت العينة من (١٨٨) فرضية صفرية ، استخدم في فحصها الإحصائي (F) وذلك في (٢٨) رسالة من رسائل الماجستير التي استخدمت تصاميم تجريبية وشبة تجريبية في مجال الإرشاد التربوي والنفسي . جمعت البيانات اللازمة لحساب مربع آيتا وقوة الاختبار الإحصائي ، كما جمعت كل البيانات الأخرى اللازمة للإجابة عن أسئلة الدراسة ، واستخرجت التكرارات والنسب المئوية للدلالة العملية والإحصائية وقوة الاختبار . وأشارت النتائج الى أن حوالي ٨٥% من الفرضيات الصفرية أُستخدم في فحصها مستوى دلالة إحصائية ٠،٠٥ وأما الباقي فأستخدم

في فحصها مستوى دلالة ٠،٠١ وأن حوالي ٦٣% من الفرضيات التي كانت دالة إحصائية لم تكن ذات دلالة عملية، أما بالنسبة لقوة الاختبار فقد اتضح أن حوالي نصف الاختبارات الإحصائية المستخدمة في فحص الفرضيات الصفرية والدالة إحصائياً تمتعت بقوة اختبار من مستوى متوسط وكبير . وأن حوالي ٢٥% من الاختبارات الإحصائية لم تكن دالة إحصائياً وكانت قوتها ضعيفة ، ومثل هذه النتيجة تثير التساؤلات حول ما إذا كان عدم التوصل الى دلالة إحصائية يعود الى ضعف قوة الاختبار أم الى عدم وجود أثر للمتغير المستقل . وأشارت النتائج أيضاً الى أن ٣،٢ %من الفرضيات الصفرية المستخدم في التحقق منها الإحصائي (F) كانت دالة عملياً ولكن لم تكن دالة إحصائياً ، مما يعني أنه كان هناك أثر كبير للمتغير المستقل ، غير أن ضعف قوة الإحصائي المستخدم حال دون الكشف عنها ، ويعود ذلك عادة الى ضعف في تصميم هذه الدراسات.

وقدم إبراهيم (١٩٩٨) دراسة تناولت الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في رسائل ماجستير علم النفس التربوي في الجامعة الأردنية .حيث هدفت الدراسة الى إظهار أهمية تضمين هذه المفاهيم في البحوث التربوية بشكل عام ومدى الجدوى من وراء استخدامها لتقييم رسائل الماجستير في مجال علم النفس التربوي /الجامعة الأردنية بشكل خاص . شملت عينة الدراسة (٣١٠) فرضية صفرية استخدم في فحصها الاختبارات الإحصائية الشائعة الاستخدام في البحوث والدراسات التربوية (χ^2, r, f, t) وذلك في (٣٣) رسالة ماجستير. وقد تم حساب كل من(η^2)،(r^2)، (ω^2)، (d) . وكان من أبرز النتائج التي توصلت إليها هذه الدراسة الى أن (٣٠،٤%) من الفرضيات الدالة إحصائياً كانت دلالاتها العملية معدومة تقريباً . وأن (٣٠،٩%) من الفرضيات الدالة إحصائياً كانت لها دلالات عملية ضعيفة ، (١١،٨%) و(٢٦،٩%) من هذه الفرضيات تمتعت بدلالات عملية من النوع المتوسط والكبير على الترتيب . بينما أقترن (١٠،٤%) من الفرضيات غير الدالة إحصائياً بدلالات عملية متوسطة وكبيرة . كما توصلت الدراسة الى أن (٥٦%) من الفرضيات غير الدالة إحصائياً كانت قوة اختباراتها الإحصائية ضعيفة .

كما أجرى الصياد (١٩٨٨م) :دراسة بعنوان " الدلالة العملية وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار (T) في البحث التربوي والنفسي ".حيث هدفت إلى التعرف على واقع الدلالة العملية من خلال استخدامات اختبار (T) في بحوث رسائل الماجستير والدكتوراه،وكذلك هدفت الدراسة إلى معرفة حجم العينة ومستوى الدلالة الشائعة في اختبار (T) ، وتكونت العينة من (١٥) دراسة في مجال علم النفس والصحة النفسية و(١١) دراسة في مجالات التربية المختلفة بالإضافة إلى (٤) رسائل ماجستير في مجال التربية ، و(٩) رسائل في مجال علم النفس، وبلغ تكرار الاختبار الإحصائي (T) ٤٦٣ مرة .وقد توصلت الدراسة إلى أن الدلالة العملية (حجم التأثير) الاختبار (T) من النوع الضعيف ، وأن الباحثين يعمدون إلى زيادة حجم العينة في سبيل الحصول على دلالة إحصائية ويميلون أيضاً لاستخدام مستويات الدلالة الإحصائية (٠,٠١ ، ٠,٠٥) حيث يتم استخدامها بعد جمع البيانات وتحليلها.

وفي دراسة أخرى قدم الصياد (١٩٨٩م) دراسة بعنوان" جدول تحديد حجم العينة في البحث السلوكي".حيث قام بإيجاد معادلات لتحديد حجم العينة لاختبارات الدلالة الإحصائية الشائعة ، وكذلك وضعها في جدول يسهل على الباحث استخدامها حسب الاختبار الإحصائي المستخدم ، وكذلك قام الباحث بحساب حجم التأثير كعامل مؤثر في حجم العينة ومتأثر به أيضاً.وتوصلت الدراسة إلى أن هناك أزمة في البحث التربوي العربي والتي تكمن في استخدام بعض الاختبارات الإحصائية بجمود دون غيرها،ودون محاولة للتفكير بما هو مناسب وما هو غير مناسب ، كما أكد على ضرورة الاهتمام بتحليل القوة الإحصائية للاختبارات.

و قدم العجلان (١٩٩٠م) دراسة بعنوان "دراسة تقويمية للأساليب الإحصائية المستخدمة في رسائل الماجستير بكلية التربية بجامعة أم القرى" وهدفت هذه الدراسة إلى التعرف على مدى ملائمة الأساليب الإحصائية في الدراسات والبحوث التي أجريت في جامعة أم القرى بكلية التربية لما ينبغي أن يكون من حيث (نوع المتغيرات، ونوع التصميم، وحجم العينة)في كل منها وتقويم هذه الأساليب في ضوء المعايير التي يجب أن تؤخذ بعين الاعتبار ،واقتراح تصور يمكن الاسترشاد

به في الاستخدام الصحيح لهذه الأساليب ،وتضمنت عينة الدراسة ١٤١ رسالة ماجستير، مـــن الرسائل التي قدمت إلى كلية التربية بجامعة أم القرى وتم التوصل إلى انه قد استخدم اختبار(ت) ١٠٦مرات، منها ٤٥ استخدام مناسب أي بنسبة ٤٢,٤٥% بينما ٦١مرة لم يكن مناسباً و٢٢ مرة يرجع إلى عدم ملائمة نوع المتغيرات للأسلوب الإحصائي المستخدم و٣٩ مرة يرجع إلى عدم ملائمة حجم العينة ونوع المتغيرات معاً .

وتوصلت هذه الدراسة إلى انه قد استخدم اختبار(كاي تربيع) ٨٢ مرة منـــها ٣١مـــرة استخدام مناسب أي بنسبة ٣٧,٨%و ٥١ مرة غير مناسب ، و٢٤ مرة يرجع السبب فيها إلى عدم ملائمة نوع المتغيرات للأسلوب الإحصائي ، بينما ٢٧ يرجع السبب إلى عدم ملائمة كل من حجـــم العينة ونوع المتغيرات معاً.

وكذلك تم استخدام اختبار (ف) ٤٩ مرة منها ١٣مرة كان الاستخدام مناسب و٣٦ مرة غير مناسب ويرجع السبب في ٣مرات إلى عدم ملائمة نوع المتغيرات للأسلوب الإحصائي وفي ٢٥ مرة إلى عدم ملائمة حجم العينة ونوع المتغيرات معاً للأسلوب الإحصائي وفي (٨) مرات إلى عدم ملائمة نوع التصميم للأسلوب الإحصائي (ف)،و وجد الباحث في هذه الدراسة أن الدلالة العملية (حجم التأثير)منخفضة في اختبارات الدلالة الإحصائية وبنسبة (٢٧،١%) من مجمـــوع اختبارات الدلالة الإحصائية المستخدمة.

وكذلك أجرى (النجار،١٤١١ه) : دراسة بعنوان: " دراسة تقويمية مقارنـــة للأســـاليب الإحصائية التي استخدمت في تحليل البيانات في رسائل الماجستير في كل من كلية التربية بجامعة أم القرى بمكة المكرمة وكلية التربية بجامعة الملك سعود بالرياض ". واستهدفت الدراسة تقويم الوضع القائم لاستخدام الأساليب الإحصائية في رسائل الماجستير في كل من كلية التربية بجامعة أم القرى وكلية التربية بجامعة الملك سعود والمقارنة بين ذلك الاستخدام ، وذلك في ضوء معايير وضـــعها الإحصائيين لذلك ، ثم محاولة التعرف على أسباب الاستخدام غير المناســـب للتغلـــب عليهـــا، بالإضافة إلى حساب الدلالة العملية لبعض الأساليب الإحصائية الشائعة الاســـتخدام في كـــلا الكليتين. واستخدم الباحث المنهج الوصفي ، حيث اختار عينة الدراسة (١٧٧) رسالة ماجستير ،

منها (١٢٧) رسالة من الرسائل التي قدمت إلى كلية التربية بجامعة أم القرى، و(٥٠) رسالة من الرسائل التي قدمت إلى كلية التربية بجامعة الملك سعود، وذلك بواقع (٥٠%) من كل مجتمع من مجتمعي الدراسة، ولحساب الدلالة العملية اختار الباحث عشوائياً (٣٥٣) اختباراً إحصائياً من كلا الكليتين وتوصل الباحث إلى النتائج التالية :

١- أن الاستخدام المناسب للأساليب الإحصائية منخفض في كلا الكليتين.

٢- أكثر الأساليب شيوعاً بكلى الكليتين هو اختبار (كاي تربيع) .

وقدم سنايدر ولاوسن (snyder and Lawson(1992) دراسة بعنوان (تقويم الدلالة الإحصائية باستخدام التقديرات الصحيحة وغير الصحيحة لحجم التأثير). هدفت إلى تناول مقاييس حجم التأثير، وكذلك تناولت الدراسة أسباب تشجيع المختصين في البحوث التربوية على استخدام حجم التأثير كمساعد تفسيري ، والأنواع المختلفة من قياسات حجم التأثير مع تصنيفها إلى مجموعات حسب مؤشرات المجتمعات والعينات ،والتحيز وعدم التحيز ،وكذلك مؤشرات نماذج التصاميم الثابتة مقابل العشوائية . توصلت الدراسة إلى أن مقاييس حجم التأثير إذا استخدمت بطريقة صحيحة فإنها تعتبر مساعدة للباحثين الذين لا يرغبون في الاعتماد فقط على اختبارات الدلالة الإحصائية في تفسير النتائج المهمة.

وأجرى نور (١٤١٣هـ) دراسة بعنوان "تقويم استخدامات اختبار كاي تربيع في رسائل الماجستير بكلية التربية بجامعة أم القرى " والتي هدفت إلى تقويم هذا الأسلوب الإحصائي في الرسائل العلمية وتحديدا رسائل الماجستير بكلية التربية بجامعة أم القرى ،والتعريف بمصادر الخطأ التي تقلل من جودة الاستخدام ،ومعرفة واقع قوة الاختبار وحجم العينة المستخدم بناء على حجم التأثير لمتغيرات الدراسة.وكان من أهم ما توصلت إليه الدراسة إن الدلالة العملية مرتفعة في اختبار كاي تربيع بنسبة (١٨ ,٧٠%) من مجموع اختبار كاي تربيع المستخدمة.كما وجدت الدراسة أن قوة الاختبار المصاحبة لاختبارات كاي تربيع مرتفعة بنسبة (٩٦%) من مجموع استخدامات اختبار كاي تربيع حيث أن مقدار القوة الإحصائية في تلك الاختبارات تجاوزت (٠,٥٠). كما

توصلت الدراسة إلى أن حجم العينة المستخدم في تلك الاختبارات كانت كبيرة ومتضخمة.

وقد قامت الصائغ (١٤١٧هـ) : بدراسة بعنوان " الدلالة الإحصائية والدلالة العملية لاختباري (ف) و(ت) دراسة تحليلية تقويمية من خلال رسائل الماجستير التي قدمت بكلية التربية بجامعة أم القرى بمكة المكرمة حتى عام ١٤١٥ هـ" . هدفت هذه الدراسة إلى التعرف على واقع الدلالات العلمية وحجم العينة المصاحبتين للدلالة الإحصائية لاختباري (ف) و(ت) في البحث التربوي والنفسي، وتكونت العينة من ١٥دراسة في مجال علم النفس والصحة النفسية و١١ دراسة في مجالات التربية المختلفة بالإضافة إلى ٤ رسائل ماجستير في مجال التربية ، و٩رسائل في مجال علم النفس.وقد كان من أهم نتائج هذه الدراسة أن واقع الدلالة العملية المصاحبة للدلالة الإحصائية كانت عند حجم تأثير كبير لاختبار (ت) ،و حجم تأثير صغير لاختبار(ف) كما أكدت الدراسة على انه ليس هناك أثر لمستوى الدلالة الإحصائية على الدلالة العملية.

وفي دراسة بلوكير (Plucker1997) عن كشف أسطورة نتيجة "عالية الدلالة":حجم التأثير في الأبحاث التربوية.حيث قامت الدراسة بدراسة ثلاث مجالات علمية ربع سنوية وأربعين مقالاً في مجلات أخرى نشرت خلال الخمس سنوات الأخيرة من الدراسة .وتوصلت الدراسة إلى أن تلك الأبحاث لا تتضمن في العادة أي مقاييس لحجوم التأثير ،كما وجد أن عبارة (عالية الدلالة) موجودة في أغلب تلك المجلات والمقالات المنشورة.وكذلك توصل الباحث إلى أهمية إدراك الباحثين لمفهوم الدلالة الإحصائية وذلك لتحسين تفسير النتائج.

وهدفت دراسة الشمراني (١٤٢١هـ) إلى التعرف على أهم مشكلات تحليل التباين وكيفية التأكد من تحققها في البيانات المراد تحليلها وتوضح البدائل المقترحة لحلها وتقويم استخدام هذا الأسلوب الإحصائي في الدراسات الجامعية.حيث أوصت الدراسة عند وجود بعض المشكلات التي تمنع تطبيق تحليل التباين كمخالفة الافتراضات ينبغي للباحث البحث بدائل أخرى مثل استخدام التحويلات المناسبة أو استخدام الإحصاء اللامعلمى، كما أوصت بأخذ عينات متساوية في الحجم للمحافظة على تجانس التباين وتجنب استخدام تحليل التباين إذا كانت أحجام العينات

في الدراسة على درجة كبيرة من التفاوت.وأظهرت نتائج الدراسة التقويمية سوء تطبيق واستخدام تحليل التباين من قبل الباحثين لأنه التطبيق يتم دون مراعاة لشروطه وافتراضاته حيث كانت نسبة مخالفة افتراض تجانس التباين في الدراسات التي تحتوي على حجوم العينات والانحراف المعيـــاري تساوي (٣٢%) أما مخالفة التوزيع الطبيعي فكانت بنسبة (٤٤%) وأن هناك قصوراً واضـــح في معرفة الباحثين بهذه المشكلات وكيفية معالجتها.

وفي دراسة لين (Lane,1999) أشارت الى وجود بعض الأخطاء الشائعة بـــشأن تفـــسير اختبار دلالة الفرض الصفري ، حيث أن هناك الكثير من الباحثين الذين قاموا بتـــضخيم نتـــائج اختبار تلك الدلالة ، وفي نفس الوقت أهملوا تماماً مؤشرات حجم التأثير، بينما الـــبعض الآخـــر منهم يرى أن الدلالة الإحصائية تعكس حجم التأثير الكلي لنتائج الدراسة . وتعتـــبر Lane أن كل من الدلالة الإحصائية وحجم التأثير مكملان لبعضهما من حيث الإشارة الى القوة الإحصائية لنتائج الدراسة .

وأجرى ثوميسون (Thompson1999) دراسة بعنوان (تحسين وتوضيح نتائج الأبحـــاث، وفائدة مؤشرات حجم التأثير كاختبارات مكملة لاختبارات الدلالة الإحصائية).حيث هـــدفت الدراسة إلى فحص مدى إشارة مجلة (الأطفال الاستثنائيين) لحجم التأثير. وقد اشـــتملت عينـــة الدراسة المقالات المنشورة من المجلد رقم (٦٣)في عام ١٦٦٩م إلى مجلد (٦٤)في عـــام ١٩٩٨م، حيث بلغ عدد المقالات المنشورة (٢٣) مقالاً. و توصل الباحث إلى أن (٢٠)مقالاً مـــن (٢٣) لم تذكر أي مؤشر لحجم التأثير.كما أشار إلى ضرورة تضمين الدراسات لمؤشرات عن حجم التأثير وفقا لتعليمات(APA) الجمعية النفسية الأمريكية.

وتناولت دراسة بابطين (٢٠٠٢): "مشكلات الدلالة الإحـــصائية في البحـــث التربـــوي وحلول بديلة " حيث هدفت الدراسة إلى الكشف عن أهم مشكلات الدلالـــة الإحـــصائية في البحث التربوي والتعرف على أهم المفاهيم والأساليب الإحصائية التي يمكن أن تقدم حلولاً لتلك المشكلات ، وصياغة أساليب جديدة لحلها والكشف عن واقع تلك المـــشكلات في الرســـائل

العلمية لكلية التربية بجامعة أم القرى.وكان أهم ما توصلت إليه الدراسة أن أهم مشكلات اختبارات الدلالة الإحصائية في البحث التربوي هي : استخدام نتائج الدلالة الإحصائية كتفسير لأهمية النتائج ، وكتفسير لتأثير المعالجة ، واستخدام قيمة الاحتمالية (P) لتقدير حجم التأثير ، وكتفسير لاحتمالية الفرضية الصفرية نفسها ، ولتقويم إعادة النتيجة ، إضافة إلى التحيز للنتائج الدالة إحصائياً ، وصياغة التساؤلات البحثية والفرضيات ، والطبيعة الثنائية لنتائج اختبار الدلالة الإحصائية ،كما توصلت الدراسة إلى أن أهم المفاهيم ولأساليب الإحصائية التي يمكن أن تقدم حلولاً لمشكلات الدلالة الإحصائية في البحث التربوي هي : تحليل القوة الإحصائية ، وتقدير حدود الثقة ، وتقديرات حجم التأثير ، وتحليل الإعادة .وكذلك أشارت الدراسة إلى واقع مشكلات الدلالة الإحصائية في رسائل الماجستير بكلية التربية بجامعة أم القرى هي على النحو التالي :

- تعتمد (٩٨%) من رسائل الماجستير بكلية التربية بجامعة أم القرى على نتائج اختبارات الدلالة الإحصائية فقط في تفسير النتائج.
- وتستخدم (٨٥،٩%) من الرسائل نتائج اختبارات الدلالة الإحصائية في تفسير تأثير المعالجات أو العوامل المستقلة .
- وتستخدم (٧،١%) من الرسائل قيم الاحتمالية (P) في تقدير حجم التأثير.
- لا يوجد أي استخدام لقيمة الاحتمالية (P) في تقدير احتمالية النتائج .
- لا يوجد أي استخدام لقيمة الاحتمالية (P) في تقدير احتمالية حدوث الفرضية الصفرية نفسها.
- نسبة النتائج الدالة إحصائياً (٤٧،٦%) من العدد الإجمالي لنتائج اختبارات الدلالة الإحصائية ولا يوجد تحيز لتلك النتائج .
- نسبة التساؤلات عن الفروق ذات الدلالة لإحصائية (٦١،٢%) من مجموع التساؤلات البحثية .

- لم تتجاوز النتائج التي قيمة الاحتمالية فيها مابين ٠،٠٥١ و ٠،٠٦ نـــسبة (٠،١%) مـــن المجموع الكلي للنتائج .

وكذلك توصلت نتائج الدراسة إلى أنه لا يوجد أي استخدام لتحليل القوة الإحصائية، ولا لتقدير حدود الثقة، ولا لتحليل الإعادة في رسائل الماجستير بجامعة أم القرى، أما مقاييس حجم التأثير فلم تستخدم إلا في ٢% من الرسائل.

أما دراسة نصار(٢٠٠٢) فقد هدفت إلى تقديم مفهوم "حجم الأثر" كأسلوب إحـــصائي مكمل لفحص الفرضيات الإحصائية. وقد تضمنت الدراسة تحليلاً رقمياً لتبرهن أن حجم الأثـــر أقل تأثراً بحجم العينة من الأساليب الإحصائية التي تستخدم لفحص الفرضيات. بالإضافة إلى ذلك قدمت الدراسة بعض المؤشرات الإحصائية التي تستخدم للدلالة على قيمة حجم الأثر في حالـــة استخدام بعض الأساليب الإحصائية وبالتحديد اختبار "ت" سواء في حالة العينات المترابطـــة أو المستقلة واختبار"ف" في حالة تحليل التباين الأحادي ومعامل ارتباط بيرسون وأخـــيراً الانحـــدار البسيط. كذلك حاولت الدراسة توضيح العلاقة بين حجم الأثر وقوة الاختبار الإحصائي حيث بينت أن حجم الأثر يرتبط إيجابياً مع قوة الاختبار المتوقع. وقدمت الدراسة نموذجاً حـــول كيفية وصف مقدار حجم الأثر في الدراسات المنشورة في المجلات العربية المتخصصة. وأظهـــرت نتائجها أنه يمكن وصف ما نسبته ٦٠% من قيم حجم الأثر والمستخرجة للأساليب الإحـــصائية المنشورة خلال خمس سنوات على أنه حجم أثر صغير أو متوسط.

وتناولت دراسة النجار (٢٠٠٥) تقويم استخدام الباحثين في الأبحاث الإدارية المنشورة في مجلة جامعة الملك سعود _العلوم الإدارية للدلالة العملية في اختبار x^2، وقد أظهرت النتـــائج أن الباحثين لم يوفقوا في الحصول على نتائج ذات دلالة إحصائية كبيرة لاختبار x^2 ، حيث بلغ عدد قيم x^2 الدالة إحصائياً (١٠٣) من مجموع قيم x^2المستخدمة في البحوث الإداريـــة المنـــشورة والبالغ عددها (٢٦١) أي نسبة (٣٩،٤٦%) ومن الملاحظ كذلك أن قيم الدلالـــة العمليـــة المنخفضة بلغت بشكل عام (١٣١) بنسبة (٥٠،١٩%) أما القيم ذات الدلالة العملية الكـــبيرة

فقد بلغت (٢٧) قيمة فقط بنسبة (١٠،٣٤%) وهذا يشير بطبيعة الحال الى أن واقـع الدلالـة العملية المصاحبة للدلالة الإحصائية لاختبار(x^2) كانت عند حجم التأثير المنخفض ، بالإضـافة الى ذلك أظهرت النتائج أن زيادة حجم العينة قد يكون له تأثير سالب على الدلالة العملية ، وأن زيادة حجم العينة في اختبار x^2 فوق (٢٠٠) أمر مكلف دون فائدة تذكر على مستوى الدلالة العملية المتحصل عليها ، لذا فأن التعرف على حجم التأثير المناسب عند إجراء الدراسة سيساعد الباحثين على تخفيض التكاليف لمجمل أبحاثهم ، وكذلك على إجـراء البحـوث والدراسـات باستخدام عينات صغيرة والحصول منها على النتائج المطلوبة .

وفي دراسة العبد القادر (١٤٢٩هـ) : بعنوان : (حجم تـأثير الاختبـارات المعلميـة واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود) . هدفت إلى التعرف على أهم الطرق الإحصائية لحساب حجم التأثير للاختبارات المعلمية واللامعلمية المـستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود للفترة ما بين (١٤١٤هـ- ١٤٢٨هـ) ، وتضمنت عينة الدراسة (١٩٢) رسالة ماجستير وذلك بواقع (٢٠%) من حجم الرسائل الصادرة بكلية التربية بجامعة الملك سعود . وتوصلت الدراسة إلى أن أكثر الأساليب الإحصائية استخداماً في كلية التربية بجامعة الملك سعود هي الأساليب المعلمية، وكذلك توصلت إلى أن حجم التـأثير للأساليب الإحصائية المستخدمة بكلية التربية تتراوح ما بين ضعيف ومتوسط كذلك أن حجـم التأثير يختلف باختلاف حجم العينة حيث أنها دالة عند مستوى (٠,٠٠١) .وأوصت الدراسـة بضرورة تضمين تدريس البدائل اللامعلمية في المقررات الدراسية وخاصة لطـلاب الدراسـات العليا، وأوصت أيضاً بضرورة تضمين تدريس طرق حساب حجم التأثير في المقررات الدراسـية وخاصة لطلاب الدراسات العليا.

التعليق على الدراسات السابقة :

من خلال استعراض الدراسات السابقة (العربية والأجنبية) تبين أن الدلالة الإحـصائية وحجم الأثر مكملان لبعضهما من حيث الإشارة الى قوة الاختبار الإحصائي لنتائج الدراسة،وأن قوة الاختبار في الأساليب الإحصائية المستخدمة كانت مابين متوسطه الى عاليه مثـل دراسـة

(تشيس وتشيس،١٩٧٦م وحجيمات وعليان،١٩٩٧م وإبـــراهيم،١٩٨٨م ولـــين،١٩٩٩م و نصار،٢٠٠٢م).أما دراسة العجلان(١٩٩٠م) فقد توصلت الى أن اغلب الاختبارات الإحصائية المستخدمة والتي تجاوزت نسبة ٥٠% كانت غير مناسبة وذلك بسبب عدم ملائمة حجم العينة ونوع المتغيرات للأسلوب الإحصائي المستخدم، كما أفادت بعض الدراسات بأن حجم العينة في البحث التربوي كبير جداً ومتضخم مثل دراسة نور (١٤١٣هـــ) والـــصياد (١٩٨٨م) .أمـــا الشمراني (١٤٢١هـــ) فقد أوصى بأخذ عينات متساوية في الحجم للمحافظة على تجانس التباين وتجنب استخدام تحليل التباين إذا كانت أحجام العينات في الدراسة على درجـــة كـــبيرة مـــن التفاوت.

كما تبين بأن قيمة مؤشرات الدلالة العملية (حجم التأثير) لاختبارات الدلالـــة الإحـــصائية كانت منخفضة وتحديداً في رسائل الماجستير في الجامعات السعودية وهذا في دراسة كلاً من العبد القادر (١٤٢٩هـــ)، وبابطين (٢٠٠٢م) ، والصائغ (١٤١٧هـــــ)، ونـــور (١٤١٣هـــــ)، والعجلان (١٩٩٠م)، والنجار (١٤١١هـــ) .وقد اتفقت معهم دراســـة الـــصياد (١٩٨٨م)، وإبراهيم (١٩٨٨م) في التوصل لنفس النتيجة التي تتضمن انخفاض قيم مؤشرات الدلالة العمليـــة وهذا يعني أن انخفاض قيمة مؤشرات الدلالة العملية موجودة ليس فقط على المستوى المحلي بل على مستوى البحث التربوي العربي .

وفي دراسة تومسيون (١٩٩٩م) يشير إلى ضعف في ذكر المؤشـــرات لحجـــم التـــأثير في المقالات المنشورة حيث أنه من عام ١٦٦٩م إلى ١٩٩٨م لم يذكر حجـــم التـــأثير إلا في (٣) مقالات فقط من ضمن (٢٠) مقالاً .ما توصل إلية أيضاً بلوكير (١٩٩٧م) حيث وجد انه خلال خمس سنوات من الدراسة لم تتضمن ثلاث مجلات علمية ربع سنوية وأربعين مقالاً أي مقـــاييس لحجم التأثير . بينما نجد دراسة سنايدر ولاوسون (١٩٩٢م) تفيد بأن مقاييس حجم التـــأثير إذا استخدمت بطريقة صحيحة فأنها تعتبر مساعدة للباحثين . وقد أوصت دراســـة العبـــد القـــادر (١٤٢٩هـــ) بضرورة تضمين تدريس طرق حساب حجم التأثير في المقررات الدراسية وخاصة لطلاب الدراسات العليا .

أما دراسة الصياد (١٩٨٩م) فقد نوهت على ضرورة الاهتمام بتحليل القوة الإحــصائية للاختبارات .وكذلك أشارة دراسة البابطين (٢٠٠٢) إلى أنه لا يوجد أي استخدام لتحليل القوة الإحصائية في عينة الدراسة التي تناولها وهذا مما يدل بشكل كبير إلى أن هناك إهمال على المستوى العربي .

واتفقت الدراسة الحالية مع الدراسات السابقة مثل دراسة (الصياد ،١٩٨٨م وحجيمــات وعليان ١٩٩٧م والصائغ وإبراهيم ،١٩٩٨م ونــصار،٢٠٠٢م والنجــار، ٢٠٠٥م وســنايدر ولاوسن،١٩٩٢م) في تناول بعض الاختبارات الإحصائية المعلمية الشائعة الاستخدام. حيــث ركزت الدراسة الحالية على اثر حجم العينة على قوة الاختبار الإحصائية وهذا ما أوصــت بــه بعض الدراسات السابقة.

الفصل الثالث

(منهج الدراسة وإجراءاتها)

أولا: منهج الدراسة

ثانيا: مجتمع الدراسة

ثالثا: عينة الدراسة

رابعا: أداة الدراسة

خامسا: صدق الأداة

سادسا: ثبات الأداة

الفصل الثالث

إجراءات الدراسة

مقدمة :

تناول هذا الفصل وصفاً للإجراءات التي تم تنفيذها من حيث تحديد منهج الدراسة، وصف مجتمع الدراسة، عينة الدراسة، بناء الأداة المناسبة ، ثم التحقق من الصدق والثبات، تحديد الأساليب الإحصائية المناسبة لتحليل بيانات الدراسة والوصول إلى النتائج .

أولاً : منهج الدراسة :

بناء على مشكلة الدراسة، وبعد الإطلاع على الدراسات السابقة، ومراجعة العديد من المناهج البحثية، حدد الباحث المنهج الملائم للدراسة الحالية وهو المنهج الوصفي والذي أشار عبيدات وآخرون (٢٠٠٥م) أنه " يعتمد على دراسة الظاهرة كما توجد في الواقع ويهتم بوصفها وصفاً دقيقاً، ويعبر عنها تعبيراً كيفياً أو كمياً، فالتعبير الكيفي يصف لنا الظاهرة ويبين خصائصها، بينما التعبير الكمي يعطينا وصفاً رقمياً لمقدار الظاهرة، أو حجمها " . وذكر العساف (٢٠٠٥م) " أن هذا المنهج لا يقتصر على جمع البيانات وتبويبها وإنما يمضي إلى ما هو أبعد من ذلك لأنه يتضمن قدراً من التفسير لهذه البيانات".

ثانياً : مجتمع الدراسة :

مجتمع الدراسة وهو يشمل " جميع المفردات أو الأشخاص الذين يكونون موضوع مشكلة البحث". (عبيدات، ٢٠٠٥م). يتكون مجتمع الدراسة الحالية من عدد من المشاهدات الإحصائية تم توليدها بواسطة برنامج (PASS11) وتراوح مجتمع الدراسة من (١٠٠) مفردة إحصائية إلى (٣٣٠) مفردة إحصائية،حسب طبيعة الأسلوب الإحصائي المستخدم. وهذه البيانات تتوافر فيها الافتراضات الخاصة بالأساليب الإحصائية المستخدمة، من حيث مستوى القياس الفئوي، التوزيع الاعتدالي.

ثالثا : عينة الدراسة :

تم اختيار عينات عشوائية من المفردات الإحصائية التي تم توليدها تراوحت من (١٠) مفردات إلى (٣٣٠) مفردة، حسب طبيعة الاختبار الإحصائي المستخدم. وروعي في الأوامر المعطاة لبرنامج (PASS11)، أن تتحقق في هذه البيانات مجموعة الافتراضات اللازمة لإجراء التحليلات الإحصائية المطلوبة، والتي تلائم البحوث التربوية سواء فيما يتعلق بالاختيار العشوائي، مستوى القياس الفئوي، التوزيع الإعتدالي، تجانس التباين.

رابعا : أداة الدراسة :

كما هو الحال في مثل الدراسة الحالية، وفي حالة دراسات المحاكاة، فإن أداة الدراسة عبارة عن جدول يتكون من مجموعة من الصفوف تمثل الحالات (الأفراد) ومجموعة من الأعمدة تمثل المتغيرات.

خامسا : صدق وثبات الأداة:

أداة الدراسة الحالية، لا تحتاج إلى حساب مؤشرات الصدق والثبات، حيث أنها ليست أداة في صورة استبيان أو اختبار، وإنما تم التأكد من طبيعة البيانات وملائمتها للاختبارات الإحصائية في الدراسة الحالية، مما يعطي مصداقية للنتائج المتحصل عليها.

سادسا : الأساليب الإحصائية :

للإجابة عن تساؤلات الدراسة تم استخدام الأساليب الإحصائية الآتية :

١. اختبار (ت) لعينة واحدة .

٢. اختبار (ت) لعينتين مترابطتين .

٣. اختبار (ت) لعينتين مستقلتين مع تساوي حجوم العينات .

٤. اختبار (ت) لعينتين مترابطتين مع عدم تساوي حجوم العينات .

٦. اختبار تحليل التباين الأحادي (ف) مع تساوي حجوم العينات .

٧. اختبار تحليل التباين الأحادي (ف) مع عدم تساوي حجوم العينات .

٨. قوة الاختبار الإحصائي .

٩. الخطأ من النوع الثاني .

الفصل الرابع

عــرض ومناقشــة النتائــج

الفصل الرابع

عرض ومناقشة النتائج

إجابة التساؤل الأول :

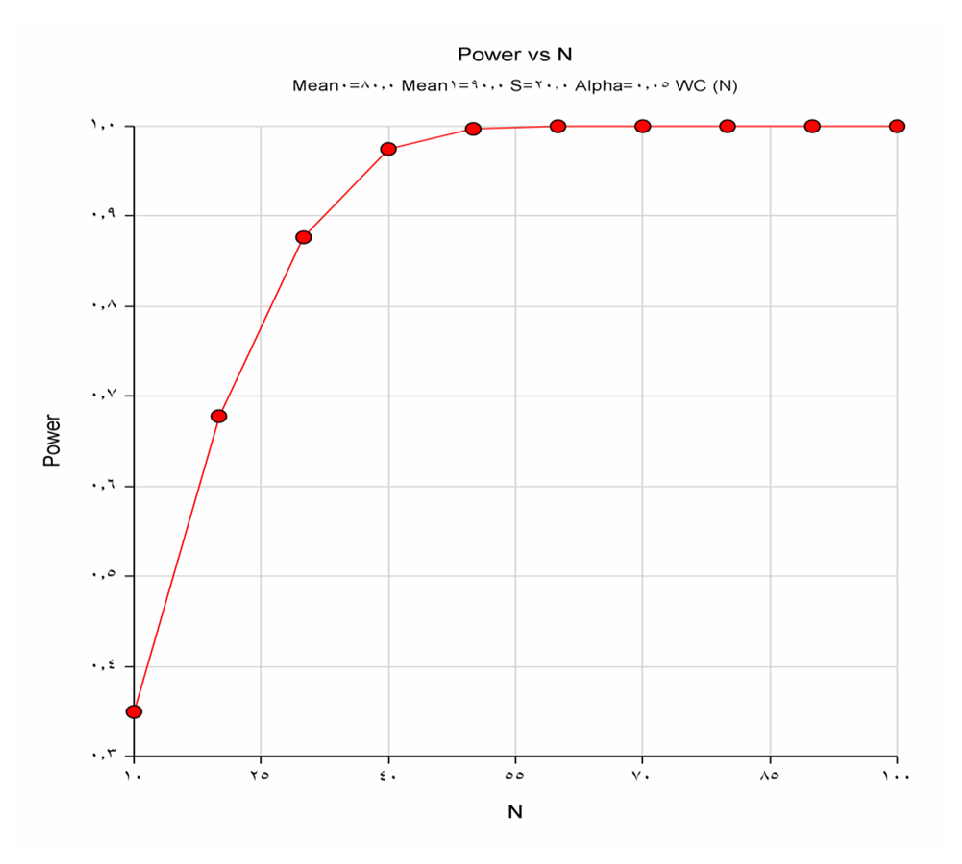
ما تأثير حجم العينة على قوة اختبار (ت) لعينة واحدة؟

للإجابة على هذا التساؤل، تم استخدام مجموعة البيانات الإحصائية المتاحة من خلال برنامج (PASS11) حيث تم توليد مجتمع من البيانات الإحصائية بحجم (N=100)، ثم أخذ عينات بأحجام من (n=10) إلى (n=100) وبمعدل زيادة (n=10)، بمتوسط عينة (M_0=80) ومتوسط فرضي (M_1=90)، وانحراف معياري (S=20)، ومجتمع البيانات الإحصائية يتوزع توزيعا معتدلا، واستخدم اختبار (ت) لعينة واحدة لاختبار الفرض الصفري بعدم وجود فروق ذات دلالة إحصائية بين متوسط العينة والمتوسط الفرضي (H_0: $\mu_0 - \mu_1 = 0$) عند مستوى (Alpha) تساوي (٠,٠٥) ضد الفرض البديل وجود فروق ذات دلالة إحصائية بين متوسط العينة والمتوسط الفرضي (H_1: $\mu_0 - \mu_1 \neq 0$)، ثم حساب قوة الاختبار الإحصائي (Power)، وفيما يلي عرض للنتائج:

جدول رقم (٤)

تأثير حجم العينة على قوة اختبار (ت) لعينة واحدة

Power	S	Mean1	Mean0	Beta	Alpha	n	Test
0.349	20	90	80	0.65	0.05	10	One-Sample T-Test Power Analysis
0.678	20	90	80	0.32	0.05	20	
0.877	20	90	80	0.12	0.05	30	
0.975	20	90	80	0.03	0.05	40	
0.997	20	90	80	0	0.05	50	
0.999	20	90	80	0	0.05	60	
1.000	20	90	80	0	0.05	70	
1.000	20	90	80	0	0.05	80	
1.000	20	90	80	0	0.05	90	
1.000	20	90	80	0	0.05	100	



شكل رقم (٢)

رسم بياني للعلاقة بين حجم العينة وقوة اختبار (ت) لعينة واحدة

تشير النتائج السابقة إلى وجود تأثير كبير لحجم العينة على قوة اختبار (ت) لعينة واحدة، حيث لوحظ أنه عندما كان حجم العينة (n=10) كانت قوة الاختبار تساوي (٠,٣٤٩) وقيمة (Beta) تساوي (٠,٦٥). وعند زيادة حجم العينة (n=20) كانت قوة الاختبار تساوي (٠,٦٧٨) وقيمة (Beta) تساوي (٠,٣٢)، وعند زيادة حجم العينة (n=30) كانت قوة الاختبار تساوي (٠,٨٧٧) وقيمة (Beta) تساوي (٠,١٢). وعند زيادة حجم العينة (n=40) كانت قوة الاختبار تساوي (٠,٩٧٤) وقيمة (Beta) تساوي (٠,٠٣)، وعندما وصل حجم العينة (n=50) اقتربت قوة الاختبار من الواحد الصحيح حيث بلغت (٠,٩٩٩) وقيمة (Beta) اقتربت من الصفر وبلغت (٠,٠٠١). ويلاحظ أنه بداية من حجم العينة (n=60) فأكثر فإن قوة الاختبار بلغت أقصى قيمة وهي (١) وانعدمت قيمة (Beta) ووصلت إلى (٠).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينة واحدة، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينة واحدة، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة (٣٠) كان مناسبا جدا للحصول على قوة اختبار عالية (٠,٨٧٧) وقيمة (Beta) منخفضة (٠,١٢).

والنتائج السابقة تتفق مع ما توصل إليه (Machin et al.,1997)

إجابة التساؤل الثاني:

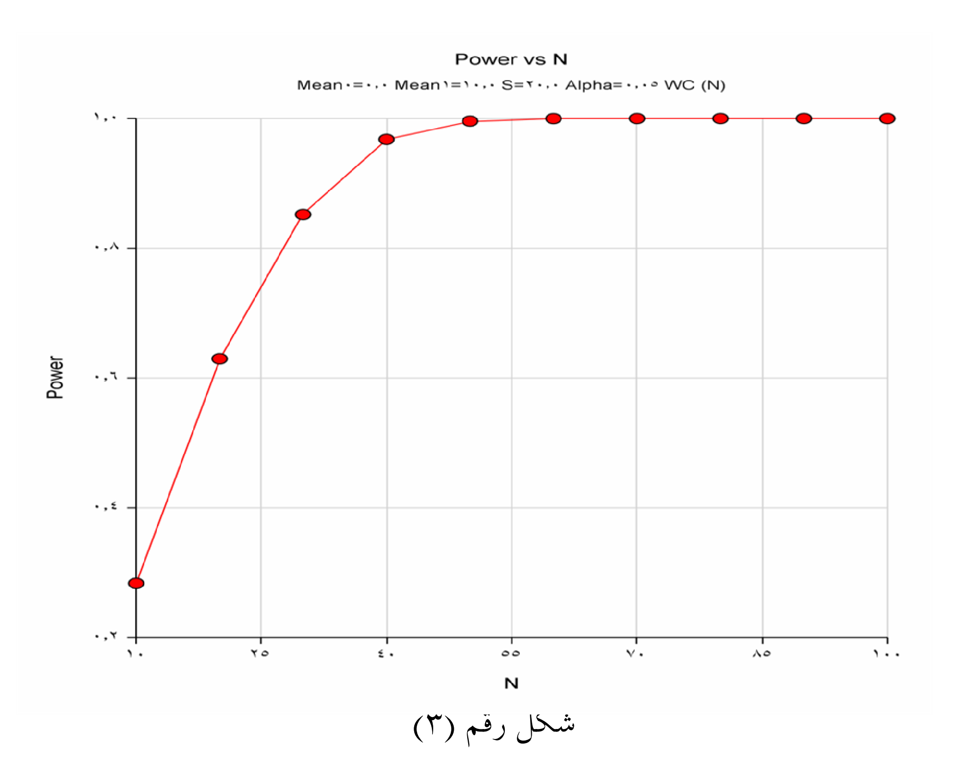
ما تأثير حجم العينة على قوة اختبار (ت) لعينتين مترابطتين؟

للإجابة على هذا التساؤل، تم استخدام مجموعة البيانات الإحصائية المتاحة من خلال برنامج (PASS11) حيث تم توليد مجتمع من البيانات الإحصائية بحجم (N=100)، ثم أخذ عينات بأحجام من (n=10) إلى (n=100) وبمعدل زيادة (n=10)، بمتوسط حسابي للعينة الأولى (M_1=80) ومتوسط حسابي للعينة الثانية (M_2=90)، وانحراف معياري (S=20)، ومجتمع البيانات الإحصائية يتوزع توزيعاً معتدلاً، وهناك استقلالية في اختيار العينات. واستخدم اختبار (ت) لعينتين مترابطتين لاختبار الفرض الصفري بعدم وجود فروق ذات دلالة إحصائية بين متوسطي العينتين ($H_0: \mu_1 - \mu_2 = 0$) عند مستوى دلالة (٠,٠٥) ضد الفرض البديل وجود فروق ذات دلالة إحصائية بين متوسطي العينتين ($H_1: \mu_1 - \mu_2 \neq 0$)، ومن ثم حساب قوة الاختبار الإحصائي (Power) وفيما يلي عرض للنتائج:

جدول رقم (٥)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مترابطتين

Power	S	$Mean_2$	$Mean_1$	Beta	Alpha	n	Test
0.284	20	90	80	0.716	0.05	10	Paired Means Power Analysis
0.630	20	90	80	0.370	0.05	20	
0.852	20	90	80	0.148	0.05	30	
0.968	20	90	80	0.032	0.05	40	
0.996	20	90	80	0.004	0.05	50	
1.000	20	90	80	0.000	0.05	60	
1.000	20	90	80	0.000	0.05	70	
1.000	20	90	80	0.000	0.05	80	
1.000	20	90	80	0.000	0.05	90	
1.000	20	90	80	0.000	0.05	100	



شكل رقم (٣)

رسم بياني للعلاقة بين حجم العينة وقوة اختبار (ت) لعينتين مترابطتين

تشير النتائج السابقة إلى وجود تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مترابطتين، حيث لوحظ أنه عندما كان حجم العينة (n=10) كانت قوة الاختبار تساوي (٠,٢٤٨) وقيمة (Beta) تساوي (٠,٧١٦). وعند زيادة حجم العينة (n=20) كانت قوة الاختبار تساوي (٠,٨٥٢) وقيمة (Beta) تساوي (٠,١٤٨)، وعند زيادة حجم العينة (n=30) كانت قوة الاختبار تساوي (٠,٨٥٢) وقيمة (Beta) تساوي (٠,١٤٨). وعند زيادة حجم العينة (n=40) كانت قوة الاختبار تساوي (٠,٩٦٨) وقيمة (Beta) تساوي (٠,٠٣٢)، وعندما وصل حجم العينة (n=50) اقتربت قوة الاختبار من الواحد الصحيح حيث بلغت (٠,٩٩٦) وقيمة (Beta) اقتربت من الصفر وبلغت (٠,٠٠٤). ويلاحظ أنه بداية من حجم العينة (n=60) فأكثر فإن قوة الاختبار بلغت أقصى قيمة وهي (١) وانعدمت قيمة (Beta) ووصلت إلى (٠).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مترابطتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مترابطتين، وفي الحالات التي تماثل

بيانات البحث الحالي فإن حجم العينة (٣٠) كان مناسباً جداً للحصول على قوة اختبار عالية (٠,٨٥٢) وقيمة (Beta) منخفضة (٠,١٤٨). والنتائج السابقة تتفق مع ما توصل إليه (Al-Sonduqchi,1990).

إجابة التساؤل الثالث:

ما تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين؟

للإجابة على هذا التساؤل، تم استخدام مجموعة البيانات الإحصائية المتاحة من خلال برنامج (PASS11) حيث تم توليد مجتمع من البيانات الإحصائية بحجم (N=200)، ثم أخذ جميع العينات الممكنة بأحجام من (n=10) إلى (n=100) وبمعدل زيادة (n=10)، بمتوسط حسابي للعينة الأولى (M_1=80) ومتوسط حسابي للعينة الثانية (M_2=90)، وانحراف معياري للعينة الأولى (S_1=10)، وانحراف معياري للعينة الثانية (S_2=20)، ومجتمع البيانات الإحصائية يتوزع توزيعاً معتدلاً، وهناك استقلالية في اختيار العينات، وتجانس للتباين. واستخدم اختبار (ت) لعينتين مستقلتين لاختبار الفرض الصفري بعدم وجود فروق ذات دلالة إحصائية بين متوسطي العينتين ($H_0: \mu_1 - \mu_2 = 0$) عند مستوى دلالة (٠,٠٥) ضد الفرض البديل وجود فروق ذات دلالة إحصائية بين متوسطي العينتين ($H_1: \mu_1 - \mu_2 \neq 0$)، ومن ثم حساب قوة الاختبار الإحصائي (Power) وفيما يلي عرض للنتائج:

أولاً: في حالة تساوي حجوم العينات ($n_1 = n_2$)

جدول رقم (٦)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين في حالة تساوي حجوم العينات

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.237	20	10	90	80	0.763	0.05	10	10	Two Ind. Sample Means Power Analysis
0.469	20	10	90	80	0.531	0.05	20	20	
0.637	20	10	90	80	0.363	0.05	30	30	
0.773	20	10	90	80	0.227	0.05	40	40	
0.856	20	10	90	80	0.144	0.05	50	50	
0.916	20	10	90	80	0.084	0.05	60	60	
0.949	20	10	90	80	0.051	0.05	70	70	
0.972	20	10	90	80	0.028	0.05	80	80	
0.983	20	10	90	80	0.017	0.05	90	90	
0.991	20	10	90	80	0.009	0.05	100	100	

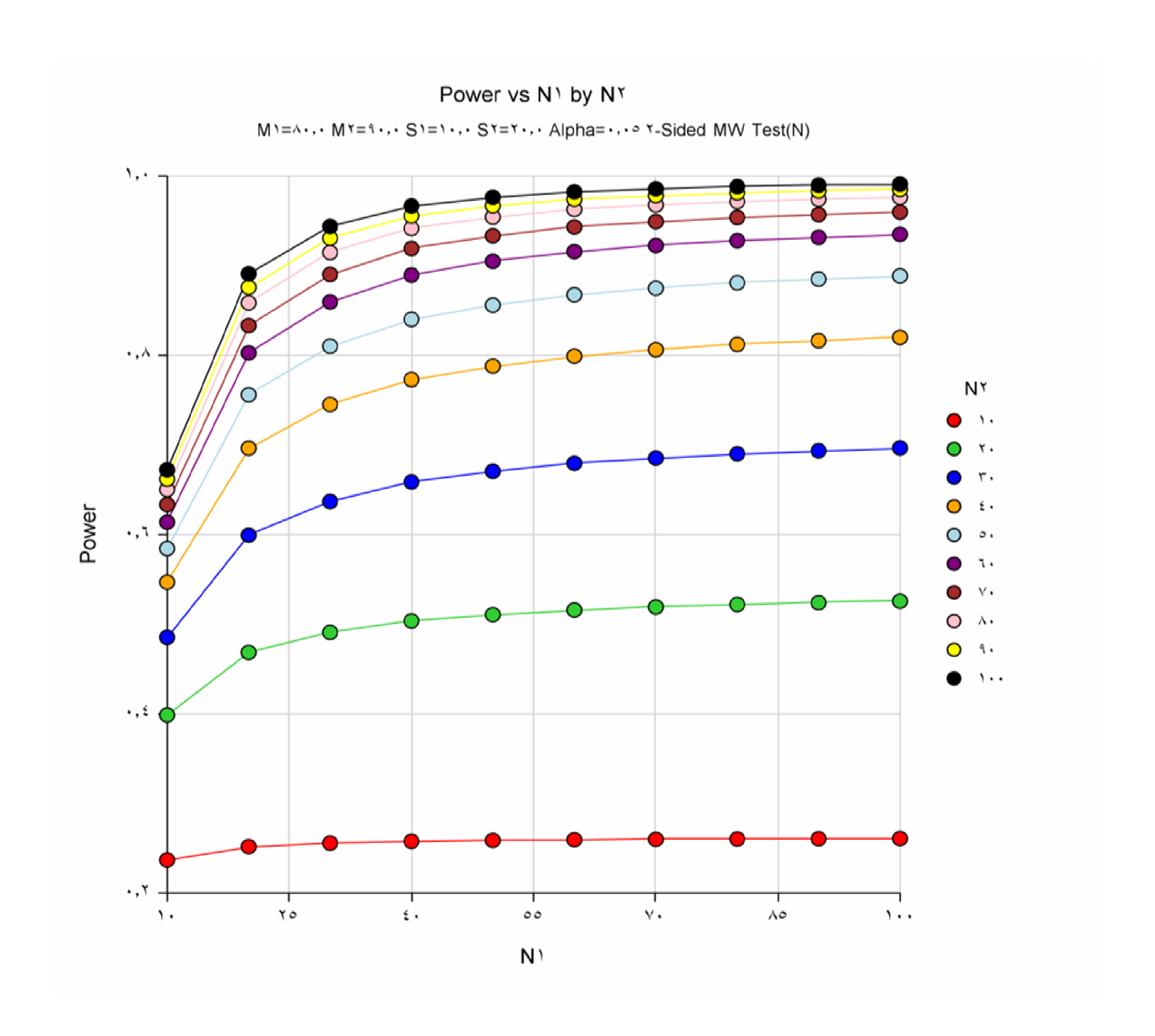
تشير النتائج السابقة إلى وجود تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، حيث لوحظ أنه عندما كان حجم العينة (n=10) كانت قوة الاختبار تساوي (٠,٢٣٧) وقيمة (Beta) تساوي (٠,٧٦٣). وعند زيادة حجم العينة (n=20) كانت قوة الاختبار تساوي (٠,٤٦٩) وقيمة (Beta) تساوي (٠,٥٣١)، وعند زيادة حجم العينة (n=30) كانت قوة الاختبار تساوي (٠,٦٣٧) وقيمة (Beta) تساوي (٠,٣٦٣). وعند زيادة حجم العينة (n=40) كانت قوة الاختبار تساوي (٠,٧٧٣) وقيمة (Beta) تساوي

(٠,٢٢٨)، وعندما وصل حجم العينة (n=50) بلغت قوة الاختبار (٠,٨٥٦) وقيمة (Beta) بلغت (٠,١١٤)، وعندما وصل حجم العينة (n=60) بلغت قوة الاختبار (٠,٩١٦) وقيمة (Beta) بلغت (٠,٠٨٤)، وعندما وصل حجم العينة (n=80) بلغت قوة الاختبار (٠,٩٧٢) وقيمة (Beta) بلغت (٠,٠٢٨)، وعندما وصل حجم العينة (n=90) بلغت قوة الاختبار (٠,٩٨٣) وقيمة (Beta) بلغت (٠,٠١٧)، وعند حجم العينة (n=100) فإن قوة الاختبار بلغت أقصى قيمة وهي (٠,٩٩١) وانخفضت قيمة (Beta) ووصلت إلى (٠,٠٠٩).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة من (٤٠ - ٥٠) كان مناسبا جدا للحصول على قوة اختبار عالية من (٠,٧٧٣ – ٠,٨٥٦) وقيمة (Beta) منخفضة (٠,٢٢٨ – ٠,١٤٤). والنتائج السابقة تتفق مع ما توصل إليه (Jerrold,1984).

ثانياً: في حالة عدم تساوي حجوم العينات ($n_1 \neq n_2$)

الحالة الأولى : ($n_1 = 10,100 \geq n_2 \geq 10$)



شكل رقم (٤)

رسم بياني لعلاقة حجم العينة بقوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات

جدول رقم (٧)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات ($n_1 = 10, 100 \geq n_2 \geq 10$)

Power	S2	S1	Mean2	Mean1	Beta	Alpha	n2	N1	Test
0.237	20	10	90	80	0.763	0.05	10	10	Two Ind. Sample Means Power Analysis
0.399	20	10	90	80	0.601	0.05	20	10	
0.485	20	10	90	80	0.515	0.05	30	10	
0.547	20	10	90	80	0.453	0.05	40	10	
0.584	20	10	90	80	0.416	0.05	50	10	
0.614	20	10	90	80	0.386	0.05	60	10	
0.634	20	10	90	80	0.366	0.05	70	10	
0.650	20	10	90	80	0.350	0.05	80	10	
0.662	20	10	90	80	0.338	0.05	90	10	
0.672	20	10	90	80	0.328	0.05	100	10	

تُشير النتائج السابقة إلى وجود تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، حيث لوحظ أنه عندما تم تثبيت حجم العينة الأولى عند ($n_1=10$) تزداد قوة الاختبار مع زيادة حجم العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) حيث كانت قوة الاختبار (٠,٢٣٧) و (٠,٦٧٢) على التوالي.

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين في حالة عدم تساوي حجوم العينات، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن النتائج تشير إلى أهمية أن حجوم العينات كبيرة في كلاً العينتين. وهذا ما أكد عليه (Machin et al.,1997)

الحالة الثانية ($n_1 = 20,100 \geq n_2 \geq 10$).

جدول رقم (٨)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات ($n_1 = 20,100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.252	20	10	90	80	0.748	0.05	10	20	Two Ind. Sample Means Power Analysis
0.469	20	10	90	80	0.531	0.05	20	20	
0.599	20	10	90	80	0.401	0.05	30	20	
0.696	20	10	90	80	0.304	0.05	40	20	
0.756	20	10	90	80	0.244	0.05	50	20	
0.803	20	10	90	80	0.197	0.05	60	20	
0.833	20	10	90	80	0.167	0.05	70	20	
0.859	20	10	90	80	0.141	0.05	80	20	
0.876	20	10	90	80	0.124	0.05	90	20	
0.891	20	10	90	80	0.109	0.05	100	20	

تشير نتائج الجدول السابق أنه بزيادة حجم العينة الأولى إلى (n_1=20) مع أحجام مختلفة من العينة الثانية بداية من (n_2=10) إلى (n_2=100) زادت قوة الاختبار من (٠,٢٥٢) إلى (٠,٨٩١) .وقيمة (Beta) انخفضت من (٠,٧٤٨) إلى (٠,١٠٩).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة الأولى (n_1=20) و (n_2=60) كان مناسبا للحصول على قوة اختبار عالية (٠,٨٠٣) وقيمة (Beta) منخفضة (٠,١٩٧).

الحالة الثالثة ($n_1 = 30,100 \ge n_2 \ge 10$):

جدول رقم (٩)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات ($n_1 = 30,100 \ge n_2 \ge 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.256	20	10	90	80	0.744	0.05	10	30	Two Ind. Sample Means Power Analysis
0.491	20	10	90	80	0.509	0.05	20	30	
0.637	20	10	90	80	0.363	0.05	30	30	
0.745	20	10	90	80	0.255	0.05	40	30	
0.810	20	10	90	80	0.190	0.05	50	30	
0.860	20	10	90	80	0.140	0.05	60	30	
0.890	20	10	90	80	0.110	0.05	70	30	
0.915	20	10	90	80	0.085	0.05	80	30	
0.931	20	10	90	80	0.069	0.05	90	30	
0.944	20	10	90	80	0.056	0.05	100	30	

تشير نتائج الجدول السابق أنه بزيادة حجم العينة الأولى إلى (n_1=30) مع أحجام مختلفة من العينة الثانية بداية من (n_2=10) إلى (n_2=100) زادت قوة الاختبار من (٠,٢٥٦) إلى (٠,٩٩٤) .وقيمة (Beta) انخفضت من (٠,٧٧٤) إلى (٠,٠٥٦).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة الأولى (n_1=30) و (n_2=50) كان مناسبا للحصول على قوة اختبار عالية (٠,810) وقيمة (Beta) منخفضة (٠,١٩٠).

الحالة الرابعة ($n_1 = 10,100 \geq n_2 \geq 10$).

جدول رقم (١٠)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات ($n_1 = 40,100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.258	20	10	90	80	0.742	0.05	10	40	Two Ind. Sample Means Power Analysis
0.504	20	10	90	80	0.496	0.05	20	40	
0.659	20	10	90	80	0.341	0.05	30	40	
0.773	20	10	90	80	0.227	0.05	40	40	
0.840	20	10	90	80	0.160	0.05	50	40	
0.890	20	10	90	80	0.110	0.05	60	40	
0.919	20	10	90	80	0.081	0.05	70	40	
0.942	20	10	90	80	0.058	0.05	80	40	
0.956	20	10	90	80	0.044	0.05	90	40	
0.967	20	10	90	80	0.033	0.05	100	40	

تشير نتائج الجدول السابق أنه بزيادة حجم العينة الأولى إلى ($n_1=40$) مع أحجام مختلفة من العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) زادت قوة الاختبار من (٠,٢٥٨) إلى (٠,٩٦٧) .وقيمة (Beta) انخفضت من (٠,٧٤٢) إلى (٠,٠٣٣).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات ($n_1=40$) و ($n_2=40$) كان مناسبا للحصول على قوة اختبار عالية (٠,٧٧٣) وقيمة (Beta) منخفضة (٠,٢٢٧).

الحالة الخامسة ($n_1 = 50,100 \geq n_2 \geq 10$).

جدول رقم (١١)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات ($n_1 = 50,100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.259	20	10	90	80	0.741	0.05	10	50	Two Ind. Sample Means Power Analysis
0.511	20	10	90	80	0.489	0.05	20	50	
0.671	20	10	90	80	0.329	0.05	30	50	
0.788	20	10	90	80	0.212	0.05	40	50	
0.856	20	10	90	80	0.144	0.05	50	50	
0.905	20	10	90	80	0.095	0.05	60	50	
0.933	20	10	90	80	0.067	0.05	70	50	
0.954	20	10	90	80	0.046	0.05	80	50	
0.967	20	10	90	80	0.033	0.05	90	50	
0.976	20	10	90	80	0.024	0.05	100	50	

تشير نتائج الجدول السابق أنه بزيادة حجم العينة الأولى إلى ($n_1=50$) مع أحجام مختلفة من العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) زادت قوة الاختبار من (٠,٢٥٩) إلى (٠,٩٧٦) .وقيمة (Beta) انخفضت من (٠,٧٤١) إلى (٠,٠٢٤).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات ($n_1=50$) و ($n_2=50$) كان مناسبا للحصول على قوة اختبار عالية (٠,٨٥٦) وقيمة (Beta) منخفضة (٠,١٤٤).

الحالة السادسة ($n_1 = 60, 100 \geq n_2 \geq 10$).

جدول رقم (١٢)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات ($n_1 = 60, 100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.260	20	10	90	80	0.740	0.05	10	60	Two Ind. Sample Means Power Analysis
0.516	20	10	90	80	0.484	0.05	20	60	
0.680	20	10	90	80	0.320	0.05	30	60	
0.799	20	10	90	80	0.201	0.05	40	60	
0.868	20	10	90	80	0.132	0.05	50	60	
0.916	20	10	90	80	0.084	0.05	60	60	
0.943	20	10	90	80	0.057	0.05	70	60	
0.963	20	10	90	80	0.037	0.05	80	60	
0.974	20	10	90	80	0.026	0.05	90	60	
0.982	20	10	90	80	0.018	0.05	100	60	

تشير نتائج الجدول السابق أنه بزيادة حجم العينة الأولى إلى ($n_1=60$) مع أحجام مختلفة من العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) زادت قوة الاختبار من (٠,٢٦٠) إلى (٠,٩٨٢) .وقيمة (Beta) انخفضت من (٠,٧٤٠) إلى (٠,٠١٨).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات ($n_1=60$) و ($n_2=40$) كان مناسبا للحصول على قوة اختبار عالية (٠,٧٩٩) وقيمة (Beta) منخفضة (٠,٢٠١).

الحالة السابعة ($n_1 = 70,100 \geq n_2 \geq 10$).

جدول رقم (١٣)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات($n_1 = 70,100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.260	20	10	90	80	0.740	0.05	10	70	Two Ind. Sample Means Power Analysis
0.519	20	10	90	80	0.481	0.05	20	70	
0.685	20	10	90	80	0.315	0.05	30	70	
0.806	20	10	90	80	0.194	0.05	40	70	
0.875	20	10	90	80	0.125	0.05	50	70	
0.923	20	10	90	80	0.077	0.05	60	70	
0.949	20	10	90	80	0.051	0.05	70	70	
0.968	20	10	90	80	0.032	0.05	80	70	
0.978	20	10	90	80	0.022	0.05	90	70	
0.986	20	10	90	80	0.014	0.05	100	70	

تشير نتائج الجدول السابق الى أنه بزيادة حجم العينة الأولى إلى ($n_1=70$) مع أحجام مختلفة من العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) زادت قوة الاختبار من (٠,٢٦٠) إلى (٠,٩٨٦) .وقيمة (Beta) انخفضت من (٠,٧٤٠) إلى (٠,٠١٤).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات ($n_1=70$) و ($n_2=40$) كان مناسبا للحصول على قوة اختبار عالية (٠,٨٠٦) وقيمة (Beta) منخفضة (٠,١٩٤).

الحالة الثامنة ($n_1 = 80,100 \geq n_2 \geq 10$).

جدول رقم (١٤)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات($n_1 = 80,100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.260	20	10	90	80	0.740	0.05	10	80	Two Ind. Sample Means Power Analysis
0.522	20	10	90	80	0.478	0.05	20	80	
0.690	20	10	90	80	0.310	0.05	30	80	
0.812	20	10	90	80	0.188	0.05	40	80	
0.881	20	10	90	80	0.119	0.05	50	80	
0.928	20	10	90	80	0.072	0.05	60	80	
0.954	20	10	90	80	0.046	0.05	70	80	
0.972	20	10	90	80	0.028	0.05	80	80	
0.981	20	10	90	80	0.019	0.05	90	80	
0.988	20	10	90	80	0.012	0.05	100	80	

تشير نتائج الجدول السابق أنه بزيادة حجم العينة الأولى إلى ($n_1=80$) مع أحجام مختلفة من العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) زادت قوة الاختبار من (٠,٢٦٠) إلى (٠,٩٨٨) .وقيمة (Beta) انخفضت من (٠,٧٤٠) إلى (٠,٠١٢).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات ($n_1=80$) و ($n_2=40$) كان مناسبا للحصول على قوة اختبار عالية (٠,٨١٢) وقيمة (Beta) منخفضة (٠,١٨٨).

الحالة التاسعة $(n_1 = 90,100 \geq n_2 \geq 10)$

جدول رقم (١٥)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات $(n_1 = 90,100 \geq n_2 \geq 10)$

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.261	20	10	90	80	0.739	0.05	10	90	Two Ind. Sample Means Power Analysis
0.524	20	10	90	80	0.476	0.05	20	90	
0.694	20	10	90	80	0.306	0.05	30	90	
0.816	20	10	90	80	0.184	0.05	40	90	
0.885	20	10	90	80	0.115	0.05	50	90	
0.932	20	10	90	80	0.068	0.05	60	90	
0.957	20	10	90	80	0.043	0.05	70	90	
0.974	20	10	90	80	0.026	0.05	80	90	
0.983	20	10	90	80	0.017	0.05	90	90	
0.990	20	10	90	80	0.010	0.05	100	90	

تشير نتائج الجدول السابق الى أنه بزيادة حجم العينة الأولى إلى $(n_1=90)$ مع أحجام مختلفة من العينة الثانية بداية من $(n_2=10)$ إلى $(n_2=100)$ زادت قوة الاختبار من (٠,٢٦١) إلى (٠,٩٩٠) .وقيمة (Beta) انخفضت من (٠,٧٣٩) إلى (٠,٠١).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات $(n_1=90)$ و $(n_2=40)$ كان مناسباً للحصول على قوة اختبار عالية (٠,٨١٦) وقيمة (Beta) منخفضة (٠,١٨٤).

الحالة العاشرة ($n_1 = 100,100 \geq n_2 \geq 10$).

جدول رقم (١٦)

تأثير حجم العينة على قوة اختبار (ت) لعينتين مستقلتين

في حالة عدم تساوي حجوم العينات($n_1 = 100,100 \geq n_2 \geq 10$)

Power	S_2	S_1	$Mean_2$	$Mean_1$	Beta	Alpha	n_2	n_1	Test
0.261	20	10	90	80	0.739	0.05	10	100	Two Ind. Sample Means Power Analysis
0.526	20	10	90	80	0.474	0.05	20	100	
0.697	20	10	90	80	0.303	0.05	30	100	
0.820	20	10	90	80	0.180	0.05	40	100	
0.889	20	10	90	80	0.111	0.05	50	100	
0.935	20	10	90	80	0.065	0.05	60	100	
0.960	20	10	90	80	0.040	0.05	70	100	
0.976	20	10	90	80	0.024	0.05	80	100	
0.985	20	10	90	80	0.015	0.05	90	100	
0.991	20	10	90	80	0.009	0.05	100	100	

تشير نتائج الجدول السابق الى أنه بزيادة حجم العينة الأولى إلى ($n_1=100$) مع أحجام مختلفة من العينة الثانية بداية من ($n_2=10$) إلى ($n_2=100$) زادت قوة الاختبار من (٠,٢٦١) إلى (٠,٩٩١) .وقيمة (Beta) انخفضت من (٠,٧٣٩) إلى (٠,٠٠٩).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ت) لعينتين مستقلتين، ومع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينات ($n_1=100$) و ($n_2=40$) كان مناسبا للحصول على قوة اختبار عالية (٠,٨٢٠) وقيمة (Beta) منخفضة (٠,١٨٠).

إجابة التساؤل الرابع

ما تأثير حجم العينة على قوة اختبار تحليل التباين الأحادي (ف) للمقارنة بين المتوسطات الحسابية؟

أولاً: حجوم العينات متساوية ($n_1= n_2 =n_3$)

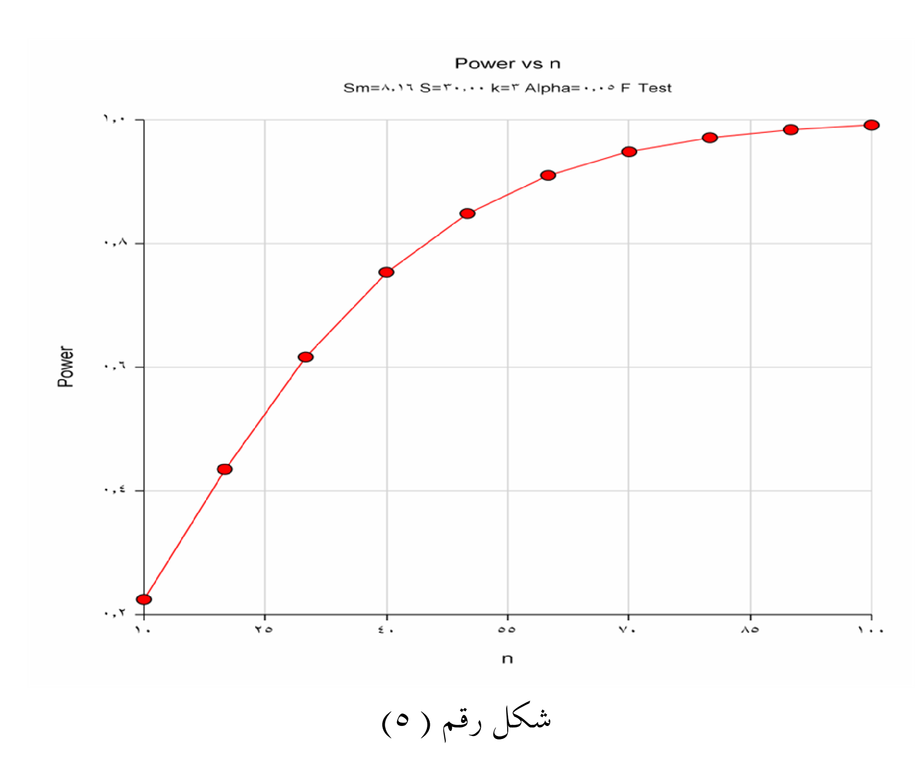
تم استخدام مجموعة البيانات الإحصائية المتاحة من خلال برنامج (PASS11) حيث تم توليد مجتمع من البيانات الإحصائية بحجم (N=300)، ثم أخذ ثلاثة عينات، كل عينة بأحجام من (n=10) إلى (n=100) وبمعدل زيادة (n=10)، بمتوسط حسابي للعينة الأولى (M_1=60) والعينة الثانية (M_2=70) والعينة الثالثة (M_2=90)، ومجتمع البيانات الإحصائية يتوزع توزيعاً معتدلاً، وهناك استقلالية في اختيار العينات، وتجانس للتباين بين المجموعات. واستخدام اختبار (ف) لاختبار الفرض الصفري بعدم وجود فروق ذات دلالة إحصائية بين متوسطات العينات الثلاث (H_0: $\mu_1 = \mu_2 = \mu_3$) عند مستوى (Alpha) تساوي (٠,٠٥) ضد الفرض البديل وجود فروق ذات دلالة إحصائية بين متوسطات العينات الثلاث (H_0: $\mu_1 \neq \mu_2 \neq \mu_3$)، ومن ثم حساب قوة الاختبار الإحصائي (Power)، وفيما يلي عرض للنتائج:

أولاً: في حالة تساوي حجوم العينات $(n_1 = n_2 = n_3)$

جدول رقم (١٧)

تأثير حجم العينة على قوة اختبار (ف) لثلاث عينات مستقلة في حالة تساوي حجوم العينات

Power	Beta	Alpha	$Mean_3$	$Mean_2$	$Mean_1$	N	n_3	n_2	n_1	Test
0.225	0.775	0.05	90	70	60	30	10	10	10	One Way ANOVA Power Analysis
0.435	0.565	0.05	90	70	60	60	20	20	20	
0.617	0.383	0.05	90	70	60	90	30	30	30	
0.754	0.246	0.05	90	70	60	120	40	40	40	
0.849	0.151	0.05	90	70	60	150	50	50	50	
0.910	0.090	0.05	90	70	60	180	60	60	60	
0.949	0.051	0.05	90	70	60	210	70	70	70	
0.971	0.029	0.05	90	70	60	240	80	80	80	
0.984	0.016	0.05	90	70	60	270	90	90	90	
0.992	0.008	0.05	90	70	60	300	100	100	100	



شكل رقم (٥)

رسم بياني للعلاقة بين حجم العينة وقوة اختبار (ف) ثلاثة عينات متساوية الحجوم

تشير النتائج السابقة إلى وجود تأثير كبير لحجم العينة على قوة اختبار (ف) للمقارنة بين مجموعة من المتوسطات الحسابية عند تساوي حجوم العينات لكل مجموعة، حيث لوحظ أنه عندما كان ($N=30$)، بواقع حجم العينة لكل مجموعة ($n=10$) كانت قوة الاختبار تساوي (٠,٢٢٥) وقيمة (Beta) تساوي (٠,٧٧٥). وعند زيادة حجم العينة ($n=20$) كانت قوة الاختبار تساوي (٠,٤٣٥) وقيمة (Beta) تساوي (٠,٥٦٥)، وعند زيادة حجم العينة ($n=30$) كانت قوة الاختبار تساوي (٠,٦١٧) وقيمة (Beta) تساوي (٠,٣٨٣). وعند زيادة حجم العينة ($n=40$) كانت قوة الاختبار تساوي (٠,٧٥٤) وقيمة (Beta) تساوي (٠,٢٤٦)، وعندما وصل حجم العينة ($n=50$) قوة الاختبار بلغت (٠,٨٤٩) وقيمة (Beta) بلغت (٠,١٥١). ويلاحظ أنه بداية من حجم العينة ($n=60$) فأكثر فإن قوة الاختبار بدأت في الزيادة عن (٠,٩) ووصلت إلى أقصى قيمة عند ($n=100$) وأصبحت (٠,٩٩٢) وقيمة (Beta) ووصلت إلى (٠,٠٠٨).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ف) للمقارنة بين مجموعة من المتوسطات الحسابية، ومع زيادة حجم العينة تزداد قوة الاختبار، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة ($40 \le n \le 50$) كان مناسبا للحصول على قوة اختبار عالية تراوحت من (٠,٧٥٤) إلى (٠,٨٤٩) وقيمة (Beta) منخفضة تراوحت من (٠,١٥١) إلى (٠,٢٤٦).

والنتائج السابقة تتفق مع ما أشار إليه (Kirk,1982 ; Fleiss,1986 and Deus,1990)

ثانيا: في حالة عدم تساوي حجوم العينات ($n_1 \neq n_2 \neq n_3$)

تم استخدام مجموعة البيانات الإحصائية المتاحة من خلال برنامج (PASS11) حيث تم توليد مجتمع من البيانات الإحصائية بحجم ($N=330$)، ثم أخذ ثلاثة عينات، كل عينة بأحجام من ($n=10$) إلى ($n=100$) وبمعدل زيادة ($n=10$)، مع مراعاة عدم تساوي حجوم العينات، وبمتوسط حسابي للعينة الأولى ($M_1=60$) والعينة الثانية ($M_2=70$) والعينة الثالثة ($M_2=90$)، ومجتمع البيانات الإحصائية يتوزع توزيعاً معتدلاً، وهناك استقلالية في اختيار العينات، وتجانس

للتباين بين المجموعات. واستخدام اختبار (ف) لاختبار الفرض الصفري بعدم وجود فروق ذات دلالة إحصائية بين متوسطات العينات الثلاث ($H_0: \mu_1 = \mu_2 = \mu_3$) عنـــد مـــستوى (Alpha) تساوي (٠,٠٥) ضد الفرض البديل وجود فروق ذات دلالة إحصائية بين متوسطات العينـــات الثلاث ($H_0: \mu_1 \neq \mu_2 \neq \mu_3$)، ومن ثم حساب قوة الاختبار الإحصائي (Power)، وفيما يلـــي عرض للنتائج:

جدول رقم (١٨)

تأثير حجم العينة على قوة اختبار (ف) لثلاث عينات مستقلة في حالة عدم تساوي حجوم العينات

Power	Beta	Alpha	M_3	M_2	M_1	N	n_3	n_2	n_1	Test
0.352	0.648	0.05	90	70	60	60	30	20	10	One Way ANOVA Power Analysis
0.537	0.463	0.05	90	70	60	90	40	30	20	
0.671	0.329	0.05	90	70	60	120	50	40	30	
0.784	0.216	0.05	90	70	60	150	60	50	40	
0.856	0.144	0.05	90	70	60	180	70	60	50	
0.914	0.086	0.05	90	70	60	210	80	70	60	
0.944	0.056	0.05	90	70	60	230	90	80	70	
0.963	0.037	0.05	90	70	60	270	100	90	80	
0.974	0.026	0.05	90	70	60	300	110	100	90	
0.985	0.015	0.05	90	70	60	330	120	110	100	

تشير النتائج السابقة إلى وجود تأثير كبير لحجم العينة على قوة اختبار (ف) للمقارنة بين مجموعة من المتوسطات الحسابية عند عدم تساوي حجوم العينات لكل مجموعة، حيث لوحظ أنه عندما كان (N=60)، كانت قوة الاختبـــار تـــساوي (٠,٣٥٢) وقيمـــة (Beta) تـــساوي (٠,٦٤٨). وعند زيادة حجم العينة (N=90) كانـــت قـــوة الاختبـــار تـــساوي (٠,٥٣٧) وقيمة (Beta) تساوي (٠,٤٦٣)، وعند زيادة حجم العينة (N=120) كانت قـــوة الاختبـــار تساوي (٠,٦٧١) وقيمة (Beta) تساوي (٠,٣٢٩)، وعند زيادة حجم العينة (N=150) كانت

قوة الاختبار تساوي (٠,٧٨٤) وقيمة (Beta) تساوي (٠,٢١٥)، وعندما وصل حجم العينة (N=180) قوة الاختبار بلغت (٠,٨٥٦) وقيمة (Beta) بلغت (٠,١٤٤). ويلاحظ أنه بداية من حجم العينة (N=210) فأكثر فإن قوة الاختبار بدأت في الزيادة عن (٠,٩) ووصلت إلى أقصى قيمة عند (N=330) وأصبحت (٠,٩٨٥) وقيمة (Beta) ووصلت إلى (٠,٠١٥).

مما سبق يمكن استنتاج أن هناك تأثير كبير لحجم العينة على قوة اختبار (ف) للمقارنة بين مجموعة من المتوسطات الحسابية مع عدم تساوي حجوم العينات، ومع زيادة حجم العينة تزداد قوة الاختبار، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة ($150 \le n \le 180$) كان مناسبا للحصول على قوة اختبار عالية تراوحت من (٠,٧٨٤) إلى (٠,٨٥٦) وقيمة (Beta) منخفضة تراوحت من (٠,٣٢٩) إلى (٠,٢١٦).

والنتائج السابقة تتفق مع ما أشار إليه (Kirk,1982 ; Fleiss,1986 and Deus,1990)

الفصل الخامس

- ملخص النتائج .
- التوصيات .
- المقترحات .

الفصل الخامس

أولاً: ملخص النتائج

- مع زيادة حجم العينة تزداد قوة اختبار (ت) لعينة واحدة، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة (٣٠) كان مناسباً جداً للحصول على قوة اختبار عالية (٠,٨٧٧) وقيمة (Beta) منخفضة (٠,١٢).
- مع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مترابطتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة (٣٠) كان مناسباً جداً للحصول على قوة اختبار عالية (٠,٨٥٢) وقيمة (Beta) منخفضة (٠,١٤٨).
- مع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة من (٤٠ - ٥٠) كان مناسباً جداً للحصول على قوة اختبار عالية من (٠,٧٧٣ – ٠,٨٥٦) وقيمة (Beta) منخفضة (٠,٢٢٨ – ٠,١٤٤).
- مع زيادة حجم العينة تزداد قوة اختبار (ت) لعينتين مستقلتين في حالة عدم تساوي حجوم العينات، وفي الحالات التي تماثل بيانات البحث الحالي فإن النتائج تشير إلى أهمية أن حجوم العينات كبيرة في كلا العينتين.
- مع زيادة حجم العينة تزداد قوة اختبار تحليل التباين الأحادي (ف) في حالة تساوي حجوم العينات، وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة ($40 \leq n \leq 50$) كان مناسباً للحصول على قوة اختبار عالية تراوحت من (٠,٧٥٤) إلى (٠,٨٤٩) وقيمة (Beta) منخفضة تراوحت من (٠,١٥١) إلى (٠,٢٤٦).
- هناك تأثير كبير لحجم العينة على قوة اختبار (ف) للمقارنة بين مجموعة من المتوسطات الحسابية مع عدم تساوي حجوم العينات، ومع زيادة حجم العينة تزداد قوة الاختبار،

وفي الحالات التي تماثل بيانات البحث الحالي فإن حجم العينة (150≤n≤80) كــان مناسباً للحصول على قوة اختبار عالية تراوحت من (٠,٧٨٤) إلى (٠,٨٥٦) وقيمــة (Beta) منخفضة تراوحت من (٠,٣٢٩) إلى (٠,٢١٦).

ثانياً: التوصيات

لما كان معلوماَ أن التوصيات تنبثق من النتائج، لذا يوصي الباحث بما يلي:

- زيادة حجم العينة إلى العدد الذي يعطي قوة اختبار مناسبة والتي تقدر بحوالي (٠,٨) في مجال العلوم التربوية، ونقص في قيمة الخطأ من النوع الثاني والذي يقدر بحوالي (٠,٢) في مجال العلوم التربوية.
- عدم المبالغة في زيادة حجم العينة عن الحد الذي يعطي قوة اختبار مناسبة.
- مراعاة أن يكون حجوم العينات مناسبة لنوع الاختبار الإحصائي المستخدم.

ثالثاً: المقترحات

يوصي الباحث بالمقترحات التالية :

- إجراء دراسة عن تأثير حجم العينة على حجم الأثر في بعض الأساليب الإحصائية.
- إجراء دراسة عن العلاقة بين حجم الأثر وقوة الاختبار.

المراجـــــع

المراجع العربية:

- إبراهيم ، ألطاف (١٩٩٨م) . واقع الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في رسائل ماجستير علم النفس التربوي في الجامعة الأردنية . رسالة ماجستير ، الأردن ، الجامعة الأردنية .
- أبو حطب ، فؤاد وصادق ، آمال (١٩٩١م). مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والتربوية والاجتماعية. القاهرة، الأنجلو المصرية.
- أبو راضي ، فتحي (١٩٩٨م) . مقدمة الطرق الإحصائية في العلوم الاجتماعية ، الطبعة الأولى ، بيروت ، دار النهضة العربية .
- أبو زيد ، مدحت (٢٠٠٢م) . الإحصاء في العلوم السلوكية الوصف الإحصائي (١) تبويب البيانات وتمثيلها بيانياً ونزعتها المركزية . الإسكندرية ، دار المعرفة الجامعية.
- أبو شعيشع ، السيد (١٩٩٧م) . الإحصاء للعلوم السلوكية . القاهرة : مكتبة النهضة المصرية .
- أبو هاشم ،السيد (٢٠٠٥) . مؤشرات التحليل البعدي لبحوث فعالية الذات في ضوء نظرية باندورا . مركز بحوث كلية التربية بجامعة الملك سعود ، العدد ٢٣٨ . الرياض .
- ابوزينه،فريد و الشايب،عبدالحافظ و عبابنة،عماد و النعيمي،محمد عبدالعال (٢٠٠٧م): مناهج البحث العلمي الإحصاء في البحث العلمي.عمان،دار المسيرة.
- آدم ، أمين إبراهيم (٢٠٠٥م) . المبادئ الأساسية الإحصائية في الطرق التطبيقية اللامعلمية، مكة المكرمة ، مكتبة الملك فهد الوطنية .
- إسماعيل ، سعيد (٢٠٠٦م) . مبادئ الإحصاء الوصفي والتطبيقي ، الأسكندرية ، مؤسسة حورس الدولية .
- الأعسر ، عبد المنعم (٢٠٠٢م) . الإحصاء لعلوم الحياة والبيئة، جدة ، الدار السعودية للنشر والتوزيع .

- بابطين ،عادل (٢٠٠٢م) . مشكلات الدلالة الإحصائية في البحث التربوي وحلول بديلة ، رسالة ماجستير ، جامعة أم القرى ،مكة المكرمة .
- باهي ، مصطفى ، وعبد العزيز، محمد ومحمد ، هيثم وسالم ، أحمد (٢٠٠٧م). الإحصاء التطبيقي باستخدام الحزم الجاهزة STATA-SPSS .
- بدر ، سالم وعابنة ، عماد (٢٠٠٧م) . مبادئ الإحصاء الوصفي والاستدلالي. عمان : دار المسيرة للنشر والتوزيع .
- برى . عدنان وهندي ، محمود وراضي ، الحسيني (١٩٩٨) . أساسيات طرق التحليل الإحصائية ،الرياض ، جامعة الملك سعود .
- البلداوي ، عبد الحميد عبد المجيد البلداوي (١٩٩٧م) . الإحصاء للعلوم الإدارية والتطبيقية، دار الشروق .
- تشاو ، لنكولن (١٩٩٠م) . الإحصاء في الإدارة . ترجمة : عبد المرضي عزام ، الرياض، دار المريخ للنشر .
- توفيق ، عبد الجبار (١٩٨٥م) . التحليل الإحصائي في البحوث التربوية والنفسية والاجتماعية – الطرق اللامعلمية . الكويت : مؤسسة الكويت لتقدم العلمي.
- حجيمات ، تحسين وعليان ، خليل (١٩٩٧) واقع الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في رسائل ماجستير الارشاد النفسي والتربوي بالجامعة الأردنية . مجلة العلوم التربوية ، المجلد (٢٤) ، العدد (٢) ، ٣٩٨- ٤٠٨ .
- الدردير ، عبد المنعم (٢٠٠٦م) . الإحصاء البارامتري واللابارامتري في اختيار فروض البحوث النفسية والتربوية والاجتماعية. القاهرة. عالم الكتب .

- الزراد ، فيصل وآخرون (١٩٨٨م) . الإحصاء النفسي والتربوي. دبي: دار القلم للنشر والتوزيع .

- سليم ،صباح وأبو حويج ،مها (٢٠٠٤م) مقدمة في الاحتمالات والإحصاء ،عمان:دار قنديل للنشر .

- سميث ، ج ملثون (١٩٨٥م) . الدليل إلى الإحصاء في التربية وعلم النفس، ترجمة: إبراهيم بسيوني عميرة، القاهرة ، دار المعارف .

- السيد ، فؤاد (٢٠٠٥م) . علم النفس الإحصائي وقياس العقل البشري ، القاهرة، دار الفكر العربي .

- الشربيني ، زكريا (٢٠٠٧) الإحصاء وتصميم التجارب في البحوث النفسية والتربوية و الإجتماعية ، القاهرة ، الأنجلو المصرية .

- الشربيني ، زكريا (٢٠٠٧م) . الإحصاء اللابارامتري في العلوم النفسية والتربوية، القاهرة ، الأنجلو المصرية .

- الشمراني،محمد (١٤٢١ه).مشكلات استخدام تحليل التباين الأحادي والمقارنات البعدية وطرق علاجها،رسالة ماجستير،مكة المكرمة،جامعة أم القرى.

- الصائغ ، ابتسام (١٤١٧هـ) . الدلالة الإحصائية والدلالة العملية لاختبار (ت) و(ف) دراسة تحليلية تقويمية من خلال رسائل الماجستير التي قدمت في كلية التربية بجامعة أم القرى بمكة المكرمة، رسالة ماجستير، مكة المكرمة، جامعة أم القرى.

- الصياد ، عبد العاطي (١٩٨٨م) . الدلالة العملية وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار(ت) في البحث التربوي والنفسي العربي – بحوث مؤتمر البحث التربوي الواقع والمستقبل ، المجلد الثاني ، القاهرة ، ص١٩٩-٢٣٠.

- الصياد،عبدالعاطي أحمد (١٩٨٩م).جدول تحديد حجم العينة في البحث السلوكي . مجلة رابطة التربية الحديثة القاهرة،العدد الأول،يناير.

- الضوي ، محسوب (٢٠٠٦م) . الإحصاء الاستدلالي المقدم في التربية وعلم النفس ، القاهرة، الانجلو المصرية .

- طه،ربيع سعيد والقاضي،ضياء(١٩٩٤م).أساسيات الإحصاء التطبيقي في المجال الزراعي،جامعة القاهرة،القاهرة،الكتاب الجامعي.

- العبد القادر،فيصل بن احمد (١٤٢٩) . حجم تأثير الاختبارات الإحصائية المعلميه واللا معلميه المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود، رسالة ماجستير ، الرياض ، كلية التربية جامعة الملك سعود .

- عبيدات، ذوقان و عبدا لحق، كايد إبراهيم و عدس، عبد الرحمن (٢٠٠٥م) : البحث العلمي مفهومه وأدواته وأساليبه، عمان،دار القلم .

- العتوم، شفيق والعارودي ، فتحي (١٩٨٦م) . الاستدلال الإحصائي وتطبيقاته في الاقتصاد والإدارة ، الجامعة الأردنية .

- العجلان ، فتحية محمد (١٩٩٠م) . دراسة تقويمية للأساليب الإحصائية المستخدمة في رسائل الماجستير بكلية التربية بجامعة أم القرى ، مكة المكرمة ، جامعة أم القرى.

- عدس ، عبد الرحمن وعبيدات ، ذوقان وعبد الحق ، كايد (٢٠٠٥ م) البحث العلمي مفهومه وأدواته وأساليبه ، الرياض دار أسامة للنشر والتوزيع .

- عدس ، عبد الرحمن (١٩٩٧م) . مبادئ الإحصاء في التربية وعلم النفس. الجزء الثاني: عمان : دار الفكر .

- العساف ، صالح (٢٠٠٦ م) . المدخل الى البحث في العلوم السلوكية ، الرياض ، العبيكان .

- علام ، صلاح الدين (٢٠٠٠) . تحليل بيانات البحوث النفسية والتربوية والاجتماعيـــة ، القاهرة ، دار الفكر العربي .
- علام ، صلاح الدين محمود (٢٠١٠م) . الأساليب الإحـــصائية الاســـتدلالي البارامتريـــة واللابارامترية في تحليل بيانات البحوث النفسية والتربوية . القاهرة: دار الفكر العربي.
- عودة ، أحمد / القاضي ، منصور (٢٠٠٢) . الإحصاء الوصفي و الاستدلالي ، الكويت، مكتبة الفلاح للنشر والتوزيع .
- عودة ، أحمد والخطيب ، أحمد (١٩٩٤م) . التحليل الإحصائي في البحوث التربوية دراسة وصفية وتحليلية.مجلة اتحاد الجامعات العربية. العدد التاسع والعشرون.
- عودة، أحمد والخليلي ، خليل (٢٠٠٠م) . الإحصاء للباحث في التربية والعلوم الإنـــسانية. الأردن : دار الأمل .
- فرح،صفوت (١٩٩٦).الإحصاء في علم النفس .القاهرة :الأنجلو المصرية .
- فهمي ، محمد (٢٠٠٥) . الإحصاء بلا معاناة المفاهيم مع التطبيقات بإستخدام برنـــامج spss ، الرياض ، مركز بحوث معهد الإدارة .
- الكناني ، ممدوح (٢٠٠٢) . الإحصاء الوصـــفي والاســـتدلالي في العلـــوم الـــسلوكية والاجتماعية . القاهرة ، دار النشر للجامعات .
- الكناني ، ممدوح (٢٠٠٢م) . الإحصاء الوصـــفي والاســـتدلالي في العلـــوم الـــسلوكية والاجتماعية . القاهرة ، دار النشر للجامعات .
- الكيلاني ، عبدالله والشريفين ، نضال (٢٠٠٧) . مدخل إلى البحث في العلوم التربويـــة والاجتماعية _ أساسياته _ مناهجه _ تـــصاميمه _ أســـاليبه الإحـــصائية. الأردن ، دار المسيرة.
- ماضي ، محمد وعثمان ، ماجد (٢٠٠٥) . الإحصاء في التربية وعلم النفس مع استخدام Spss و Minitab ، الطبعة الثانية ، دبي ، دار القلم للنشر والتوزيع .

- ماضي ، محمد وعثمان ، ماجد (٢٠٠٥م) . الإحصاء في التربية وعلم النفس مع استخدام Minitab, SPSS الطبعة الثانية، دبي، دار القلم للنشر والتوزيع.
- المالكي، مرضي راضي (١٤٢٢هـ). واقع استخدام الأساليب الإحصائية في أبحاث التربية الإسلامية في بعض الجامعات السعودية. رسالة دكتوراه غير منشورة. مكة المكرمة: كلية التربية جامعة أم القرى.
- منصور ، رشدي فام (١٩٩٧) . حجم التأثير _ الوجه المكمل للدلالة الإحصائية ، المجلة المصرية للدراسات النفسية ، العدد (١٦) ، ٥٧- ٧٥ .
- المنيزل ، عبدالله (٢٠٠٠) . الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الإحصائية Spss ، عمان ، دار وائل للنشر .
- النبهان ، موسى (٢٠٠١) . أساسيات الإحصاء في التربية والعلوم الإنسانية والتربوية ، الكويت ، مكتبة الفلاح للنشر والتوزيع .
- النجار ، عبدالله (٢٠٠٥) . الدلالة الإحصائية والدلالة العلمية لاختبار كا٢ في البحوث الإدارية المنشورة (دراسة تقويميه) . المجلة العربية للعلوم الإدارية .(١٢) ، (٢) ، ١٦٩- ١٩٣ .
- النجار، عبد الله عمر (١٤١١هـ) . دراسة تقويمية للأساليب الإحصائية التي استخدمت في تحليل البيانات في رسائل الماجستير في كل من كلية التربية بجامعة أم القرى بمكة المكرمة وكلية التربية بجامعة الملك سعود بالرياض، رسالة ماجستير غير منشورة ، بمكة المكرمة ، كلية التربية ، جامعة أم القرى .
- نصار، يحي (٢٠٠٢). حجم الأثر كأسلوب إحصائي مكمل لفحص الفرضيات الإحصائية، مركز بحوث كلية التربية – الرياض جامعة الملك سعود، العدد ١٧٦.
- النعيمي ، محمد والبياتي ، حسين (٢٠٠٦) الإحصاء المتقدم في العلوم التربوية والتربية البدنية مع تطبيقات SPSS ، الوراق للنشر والتوزيع .
- نور ، رجاء محمد (١٤١٣هـ) . تقويم استخدامات اختبار كاي تربيع في رسائل الماجستير

بكلية التربية – جامعة أم القرى. رسالة ماجستير غير منشورة . مكة المكرمة: كلية التربية ، جامعة أم القرى .

- هويل ، بول جي (١٩٨٤م) . المبادئ الأولية في الإحصاء . ترجمة: بدرية عبد الوهاب ومحمد الشربيني ، الطبعة الرابعة ، دار وايلي وأبنائه .
- الهيتي ، صلاح الدين (٢٠٠٤) . الأساليب الإحصائية في العلوم الإدارية : تطبيقات بأستخدام SPSS ، عمان ، دار للطابعة والنشر .
- هيكل ، عبد العزيز فهمي (١٤٠٦هـ) . موسوعة المصطلحات الاقتصادية والإحصائية. الطبعة الثانية ، دار النهضة العربية ، بيروت – لبنان .

المراجع الأجنبية:

- Al-Sounduqchi , 1990. Power of statistical test. www.pass11.com.
- American psychological Association (2001) . publication Manual of the American Psychological Association (5th ed .) Washington ,DC: Author.
- Anderson, D, Burnham , K ,and Thompson, W (2000) . Null Hypothesis Testing : problems , prevalence , and an Alternative . Journal of Wildlife Management , 64 , 912-923.
- Arnemnn, K.(2003). A Review of he panoply of effectsize choices.(Eric Document Reproductive Service No. ED 473811).
- Aron, A., & Aron, E.N. (1997). Statistics for the behavioral and social sciences: A brief course. NJ: Prentice Hall
- Bird, K (2002) . Confidence Intervals for Effect sizes in Anlysis of Variance . Educational and psychological Measurement , 62, 2, 197-226.
- Brewer, J (1972) . On the power of Statistical Tests . American Educational Research Journal, 9, (3) . 391-401 . Chase, L. and chase , R

(1976) A Statistical power Analysis of Applied Psychological Research , Journal of Applied Psychology, 61: 234-237 .

- Chase, L. and chase , R (1976) A Statistical power Analysis of Applied Psychological Research , Journal of Applied Psychology, 61: 234-237 .
- Cohen, J. (1977). Statistical power analysis for the behavioral sciences , New York: Academic Press.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.
- Craig, J.R., Eison , C.L., & Metz , L.P. (1976). Significance tests and their interpretation: An example utilizing published research and Omega Squared. Bulletin of the Psychonomic society , 7, 280-282.
- Deus,1990. Power of statistical test. www.pass11.com.
- Fan, Xitao.(2001).Statistical significance and effect size in education research:Two sides of a coin. Journal of Educational Reseach,v94, n5, 257-282.
- Fern, E.,Monroe,K.B. (1996).Effect-size estimates: Issues and problems in interpretation. Journal of Consumer Research, 23, 89-104.
- Filder, F and Thompson, B (2001) .Computing Correct Confidence Intervals for ANOVA fixed- and Random Effect Sizes. Educational and psychological Measurement, 61, ,4,575-604. 12- Galarza, A(1993). What is the Probility of Rejecting, the Null Hypothesis ? Statistical Power in Research . Paper Presented at the Annual meeting of the mid-south Educational Research Assoceation, 22nd – new Orleans . ERIC Document Reprodutive Service, NO. ED 364593.
- Fleiss,1986 . Power of statistical test. www.pass11.com.

- Haase. R, waechter, D. and Solomon,G. (1982). How Significant is a Significant Difference ? Average Effect Sizes of Research in Counseling Psychology, Journal of Counseling Psychology, 29,58-65.
- Haase. R, waechter, D. and Solomon,G. (1982). How Significant is a Significant Difference ? Average Effect Sizes of Research in Counseling Psychology, Journal of Counseling Psychology, 29,58-65.
- Hays, W.L. (1963). Statistics for Psychologists. New York: Holt , Rinehart , and Winston.
- Howell, D.C. (1995). Fundamental statistics for the behavioral sciences (3rd ed.). CA: Belmont.
- Huberty, C (2002) . A History of Effect Size Indices. Educational and Psychological Measurement, 62 , 2, 227-240.
- Huston, H (1993). Meaning fullness, Statistical Significance, Effect Size and power Analysis: A General Discussion with Implications for MANOVA, Paper Presented at the annual meeting of the mid-south Educational Research Association, 22nd , New Orleans, ERIC Document Reprcuctive Service . NO ED 364608.
- Kirk, 1982. Power of statistical test. www.pass11.com.
- Lane, G (1999). Show my the Magnitude! The Consequences of Overemphasis on Null Hypothesis Significance Testing. Paper Presented at the annual meeting of the mid-south Educational Research Association (al, November 16-19, 1999). (Eric document Reprodutive service no . ED 436557).
- Lane, G.(1999).Show my the magnitude! the consequences of overemphasis on null hypothesis significance testing. Paper presented at the annual meeting of the Mid-South Educational Research

Association(Al November 16-19,1999).(Eric Document Reproductive Service No. ED 436557).

- Machin et al., 1997. Power of statistical test. www.pass11.com.
- Maxwell, S and Delaney, H (2004). Designing experiments and analyzing data : A model comparisonperspective (2 nd ed) . Mahwah , NJ : Lawrence Erlbaum.
- Parker , R and Hagan-Burke, S (2007). Useful Effect Size Interpretations for Single Case Research . Behavior Therapy, 83 , 95-105.
- Pluker,Jontahan A.(1997).Debunking The Myth of ((Highly significant)) Result :Effult size in Gifted Education Research. Roeper Review. V20 n2, pp122-26,1997.
- Snyder, P. and Lawson , S (1992) . Evaluating Significance using Corrected and Uncorrected Magnifude of Effect Size Estimates. Paper Presented at Annual Meting of the Americon Experimental Research Association (San Francisco,CA,Aril 20-24,1992).
- Steinberg, L and Thissen, D (2006). Using Effect Sizes for Research Reporting : Examples Using Item Response Theory to Analyze Differential Item Functioning . Psychological Methods , 11, 4, 402-415.
- Thompson, Bruce.(1999). Improving Pesearch Clarity and Usefulnass with Effect size Indices as Aupplement to Statistical significance Tests.Exceptional Chidren V65, n3,pp.329-337,1999.

الملاحق



بسم الله الرحمن الرحيم

استمارة تعديل (عنوان موضوع) رسالة علمية

لطلبة الدراسات العليا بجامعة أم القرى



عنوان موضوع الرسالة المسجل : اثر حجم العينة وعدد المتغيرات المستقلة على قوة اختبار (ف) في التصاميم العاملية .

عنوان موضوع الرسالة بعد التعديل : تأثير حجم العينة وعدد المتغيرات على قوة الاختبار الاحصائي

الطالب / ~~الطالبة~~: محمد ابراهيم احمد الشاردي الرقم الجامعي ٤٢٩٨٠٤٥٦

الدرجة العلمية: ☐ دبلوم عالي ☑ ماجستير ☐ دكتوراه

الكلية: التربية القسم: علم النفس التخصص: احصاء وبحوث

أوافق على تعديل عنوان موضوع الرسالة وفق ماورد أعلاه .

رئيس القسم

المشرف على الرسالة

الاسم : أ.د ربيع سعيد طه

التوقيع :

التاريخ : ٢٢/١٠/١٤٣٢هـ

أتقدم أنا الطالب / ~~الطالبة~~ : محمد ابراهيم احمد الشاردي ، إلى معهد البحوث العلمية وإحياء التراث الإسلامي، وأرجو اتخاذ الإجراءات اللازمة لتعديل عنوان الموضوع أعلاه ، حسب المتبع في هذا الشأن، مع جزيل الشكر والتقدير .

التوقيع :

التاريخ : ٢٢/١٠/١٤٣٢هـ

بعد البحث النصي في قواعد البيانات المتوفرة لدى المعهد ، بشأن عنوان موضوع الرسالة العلمية المشار إلى بياناتها أعلاه ، انتهينا إلى ما يأتي :

☑ تم تعديل عنوان موضوع الرسالة ، وحذف العنوان المسجل

☐ لم يتم تعديل عنوان الرسالة ، للأسباب التالية : ....

الختم

الموظف المختص بالمعهد

الاسم :

التوقيع :

لم تتم دراسته

الرقم / ٤٩٦٥ التاريخ / ٢٢/١٠/١٤٣٢